

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 101.709 seconds

(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-13

Perfect score: 73

Sequence: 1 IEGETLRQWLARA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A_Geneseq_21.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	14	AAW09463	Aaw09463 Thrombopo
2	73	100.0	14	AAW09468	Aaw09468 Thrombopo
3	73	100.0	14	AAW33030	Aaw33030 Thrombopo
4	73	100.0	14	AAW33034	Aaw33034 Thrombopo
5	73	100.0	14	AAW36774	Aaw36774 Thrombopo
6	73	100.0	14	ADI24843	Adi24843 AF 12505
7	73	100.0	14	AAI96515	Aay96515 Thrombopo
8	73	100.0	14	AAI96512	Aabi9652 TPO-mimet
9	73	100.0	14	AAU25827	Aau25827 Human thr
10	73	100.0	14	AAU26004	Aau26004 Human thr
11	73	100.0	14	ABB72853	Abb72853 TPO mimet
12	73	100.0	14	ABP51669	Abp51669 Thrombopo
13	73	100.0	14	AAE18011	Aae18011 Human lig
14	73	100.0	14	ABG71747	Abg71747 TPO recep
15	73	100.0	14	ABR62907	Abrc62907 thrombopo
16	73	100.0	14	ADC33697	Adc33697 Erythrope
17	73	100.0	14	ADN59652	Adn59652 Thrombopo
18	73	100.0	14	ADL27293	Adl27293 Amino aci
19	73	100.0	14	ADM72483	Adm72483 TPO mimet
20	73	100.0	14	ADQ16584	Adq16584 Agonist T
21	73	100.0	14	ADQ16582	Adt92482 Thrombopo
22	73	100.0	14	ADU70210	Adu70210 Thrombopo
23	73	100.0	14	ADU75982	Adu75982 Peptide-b
24	73	100.0	14	ADV44319	Adv44319 Agonist T

25	73	100.0	14	AEI12792	Aebi2792 TPO mimet
26	73	100.0	15	AAW35416	Aaw35416 Thrombopo
27	73	100.0	15	AAW36776	Aaw36776 Thrombopo
28	73	100.0	15	AAW66712	Aaw66712 Peptide c
29	73	100.0	15	AAW20684	Aaw20684 Thrombocy
30	73	100.0	15	AAU25996	Aau25996 Human thr
31	73	100.0	15	AAU25831	Aau25831 Human thr
32	73	100.0	15	ABP51670	Abp51670 Thrombopo
33	73	100.0	15	ABR62908	Abrc62908 Thrombopo
34	73	100.0	15	ADM72485	Adm72485 TPO mimet
35	73	100.0	15	ADM72479	Adm72479 TPO mimet
36	73	100.0	15	ADM72478	Adm72478 TPO mimet
37	73	100.0	15	ADM72533	Adm72533 TPO mimet
38	73	100.0	15	ADM72522	Adm72522 TPO mimet
39	73	100.0	15	ADM72523	Adm72523 TPO mimet
40	73	100.0	15	ADM72482	Adm72482 TPO mimet
41	73	100.0	15	ADQ16585	Adq16585 TPO mimet
42	73	100.0	15	ADT92483	Adt92483 Modified
43	73	100.0	15	ADU70209	Adu70209 Thrombopo
44	73	100.0	15	ADU75981	Adu75981 Peptide-b
45	73	100.0	15	ADV44320	Adv44320 Agonist T
46	73	100.0	15	AEI12793	Aebi2793 TPO mimet
47	73	100.0	16	AAW19534	Aaw19534 Thrombopo
48	73	100.0	16	AAW33035	Aaw33035 Thrombopo
49	73	100.0	16	AAW36775	Aaw36775 Thrombopo
50	73	100.0	16	AAW36771	Aaw36771 Thrombopo
51	73	100.0	16	AAW66709	Aaw66709 Peptide c
52	73	100.0	16	AAW66713	Aaw66713 Peptide c
53	73	100.0	16	AAW66733	Aaw66733 Peptide c
54	73	100.0	16	AAW66716	Aaw66716 Peptide c
55	73	100.0	16	AAU26005	Aau26005 Human thr
56	73	100.0	16	AAU26043	Aau26043 Human thr
57	73	100.0	16	AAU25832	Aau25832 TPO mimet
58	73	100.0	16	ADM72532	Adm72532 TPO mimet
59	73	100.0	16	ADM72484	Adm72484 TPO mimet
60	73	100.0	18	AAI16957	Aai16957 PEGylated
61	73	100.0	18	AAI16956	Aai16956 PEGylated
62	73	100.0	18	ABP51687	Abp51687 TPO mimet
63	73	100.0	18	ABP51689	Abp51689 TPO mimet
64	73	100.0	18	ABP51688	Abp51688 TPO mimet
65	73	100.0	18	ABP51677	Abp51677 TPO mimet
66	73	100.0	18	ABP51686	Abp51686 TPO mimet
67	73	100.0	18	ABP51674	Abp51674 TPO mimet
68	73	100.0	18	ABP51693	Abp51693 TPO mimet
69	73	100.0	18	ABP51684	Abp51684 TPO mimet
70	73	100.0	18	ABP51683	Abp51683 TPO mimet
71	73	100.0	18	ABP51685	Abp51685 TPO mimet
72	73	100.0	18	ABP51691	Abp51691 TPO mimet
73	73	100.0	18	ABP51673	Abp51673 TPO mimet
74	73	100.0	18	ABP51690	Abp51690 TPO mimet
75	73	100.0	18	ABP51675	Abp51675 TPO mimet
76	73	100.0	18	ABP51692	Abp51692 TPO mimet
77	73	100.0	18	ADN59812	Adn59812 Thrombopo
78	73	100.0	18	ADQ16611	Adq16611 TPO mimet
79	73	100.0	18	ADQ16619	Adq16619 TPO mimet
80	73	100.0	18	ADQ16621	Adq16621 TPO mimet
81	73	100.0	18	ADQ16641	Adq16641 TPO mimet
82	73	100.0	18	ADQ16646	Adq16646 TPO mimet
83	73	100.0	18	ADQ16607	Adq16607 TPO mimet
84	73	100.0	18	ADQ16615	Adq16615 TPO mimet
85	73	100.0	18	ADQ16627	Adq16627 TPO mimet
86	73	100.0	18	ADQ16625	Adq16625 TPO mimet
87	73	100.0	18	ADQ16617	Adq16617 TPO mimet
88	73	100.0	18	ADQ16629	Adq16629 TPO mimet
89	73	100.0	18	ADQ16613	Adq16613 TPO mimet
90	73	100.0	18	ADQ16623	Adq16623 TPO mimet
91	73	100.0	18	ADQ16605	Adq16605 TPO mimet
92	73	100.0	18	ADQ16609	Adq16609 TPO mimet
93	73	100.0	18	ADQ16608	Adq16608 TPO mimet
94	73	100.0	18	ADV44355	Adv44355 Agonist T
95	73	100.0	18	ADV44345	Adv44345 Agonist T
96	73	100.0	18	ADV44351	Adv44351 Agonist T
97	73	100.0	18	ADV44357	Adv44357 Agonist T
98	73	100.0	18	ADV44349	Adv44349 Agonist T

98	73	100.0	18	9	ADV44343	Adv44343 Agonist T	171	73	100.0	29	3	AAB16975	Aab16975 TPO-mimet
99	73	100.0	18	9	ADV44347	Adv44347 Agonist T	172	73	100.0	29	3	AAB16976	Aab16976 TPO-mimet
100	73	100.0	18	9	ADV44384	Adv44384 Agonist T	173	73	100.0	29	3	AAB17286	Aab17286 TPO-mimet
101	73	100.0	18	9	ADV44353	Adv44353 Agonist T	174	73	100.0	29	3	AAB16970	Aab16970 TPO-mimet
102	73	100.0	18	9	ADV44363	Adv44363 Agonist T	175	73	100.0	29	5	ABB72862	Abb72862 TPO mimet
103	73	100.0	18	9	ADV44361	Adv44361 Agonist T	176	73	100.0	29	5	ABB72861	Abb72861 TPO mimet
104	73	100.0	18	9	ADV44379	Adv44379 Agonist T	177	73	100.0	29	5	ABB72856	Abb72856 TPO mimet
105	73	100.0	18	9	ADV44344	Adv44344 Agonist T	178	73	100.0	29	7	ADJ73011	Adj73011 TPO mimet
106	73	100.0	18	9	ADV44359	Adv44359 Agonist T	179	73	100.0	29	7	ADJ73006	Adj73006 TPO mimet
107	73	100.0	18	9	ADV44365	Adv44365 Agonist T	180	73	100.0	29	8	ADJ52646	Adj52646 CHI Delet
108	73	100.0	18	9	ADV44367	Adv44367 Agonist T	181	73	100.0	29	8	ADJ52641	Adj52641 CHI Delet
109	73	100.0	18	9	ADV44468	Adv44468 Agonist T	182	73	100.0	29	8	ADJ51602	Adj51602 CHI Delet
110	73	100.0	18	9	ABE12826	Aeb12826 TPO mimet	183	73	100.0	29	8	ADJ51607	Adj51607 CHI Delet
111	73	100.0	18	9	ABE12820	Aeb12820 TPO mimet	184	73	100.0	30	3	AAB17287	Aab17287 TPO-mimet
112	73	100.0	18	9	ABE12840	Aeb12840 TPO mimet	185	73	100.0	31	3	AAB17288	Aab17288 TPO-mimet
113	73	100.0	18	9	ABE12857	Aeb12857 Antibody	186	73	100.0	31	3	AAB16974	Aab16974 TPO-mimet
114	73	100.0	18	9	ABE12834	Aeb12834 TPO mimet	187	73	100.0	31	3	AAB16973	Aab16973 TPO-mimet
115	73	100.0	18	9	ABE12830	Aeb12830 TPO mimet	188	73	100.0	31	5	ABB72860	Abb72860 TPO mimet
116	73	100.0	18	9	ABE12939	Aeb12939 TPO mimet	189	73	100.0	31	5	ABB72859	Abb72859 TPO mimet
117	73	100.0	18	9	ABE12832	Aeb12832 TPO mimet	190	73	100.0	31	7	ADJ73009	Adj73009 TPO mimet
118	73	100.0	18	9	ABE12852	Aeb12852 TPO mimet	191	73	100.0	31	7	ADJ73010	Adj73010 TPO mimet
119	73	100.0	18	9	ABE12816	Aeb12816 TPO mimet	192	73	100.0	31	8	ADJ52644	Adj52644 CHI Delet
120	73	100.0	18	9	ABE12818	Aeb12818 TPO mimet	193	73	100.0	31	8	ADJ52645	Adj52645 CHI Delet
121	73	100.0	18	9	ABE12828	Aeb12828 TPO mimet	194	73	100.0	31	8	ADJ51606	Adj51606 CHI Delet
122	73	100.0	18	9	ABE12836	Aeb12836 TPO mimet	195	73	100.0	31	8	ADJ51605	Adj51605 CHI Delet
123	73	100.0	18	9	ABE12822	Aeb12822 TPO mimet	196	73	100.0	32	3	AAV96520	Aay96520 Thrombopo
124	73	100.0	18	9	ABE12824	Aeb12824 TPO mimet	197	73	100.0	32	3	AAB17289	Aab17289 TPO-mimet
125	73	100.0	18	9	ABE12838	Aeb12838 TPO mimet	198	73	100.0	32	3	AAB17297	Aab17297 TPO-mimet
126	73	100.0	19	5	ABB73391	Abb73391 TPO-mimet	199	73	100.0	33	3	AAB17290	Aab17290 TPO-mimet
127	73	100.0	19	5	ABB73390	Abb73390 TPO-mimet	200	73	100.0	34	3	AAV96527	Aay96527 Thrombopo
128	73	100.0	20	3	AAB18003	Aab18003 FC-TMP pe	201	73	100.0	34	3	AAB17291	Aab17291 TPO-mimet
129	73	100.0	20	3	AAB17929	Aab17929 TPO-mimet	202	73	100.0	35	3	AAB17292	Aab17292 TPO-mimet
130	73	100.0	20	5	ABB73403	Abb73403 TPO mimet	203	73	100.0	36	3	AAV96525	Aay96525 Thrombopo
131	73	100.0	21	7	ADN59687	Adn59687 Thrombopo	204	73	100.0	36	3	AAV96523	Aay96523 Thrombopo
132	73	100.0	22	7	ADN59819	Adn59819 TMP pepti	205	73	100.0	36	3	AAV96524	Aay96524 Thrombopo
133	73	100.0	22	8	ADQ16714	Adq16714 Immunoglo	206	73	100.0	36	3	AAV96526	Aay96526 Thrombopo
134	73	100.0	22	8	ADQ16713	Adq16713 Immunoglo	207	73	100.0	36	3	AAB17307	Aab17307 TPO-mimet
135	73	100.0	22	8	ADQ16709	Adq16709 Immunoglo	208	73	100.0	36	3	AAB17293	Aab17293 TPO-mimet
136	73	100.0	22	8	ADQ16706	Adq16706 Immunoglo	209	73	100.0	36	3	AAB17303	Aab17303 TPO-mimet
137	73	100.0	22	8	ADQ16699	Adq16699 TPO mimet	210	73	100.0	36	3	AAB16963	Aab16963 TPO-mimet
138	73	100.0	22	8	ADQ16712	Adq16712 Immunoglo	211	73	100.0	36	3	AAB17301	Aab17301 TPO-mimet
139	73	100.0	22	8	ADQ16707	Adq16707 Immunoglo	212	73	100.0	36	3	AAB17306	Aab17306 TPO-mimet
140	73	100.0	22	8	ADQ16711	Adq16711 Immunoglo	213	73	100.0	36	5	ABB72403	Abb72403 TPO-mimet
141	73	100.0	22	8	ADQ16708	Adq16708 Immunoglo	214	73	100.0	37	3	AAB17294	Aab17294 TPO-mimet
142	73	100.0	22	8	ADQ16710	Adq16710 Immunoglo	215	73	100.0	38	3	AAB17295	Aab17295 TPO-mimet
143	73	100.0	22	9	ADV44435	Adv44435 Modified	216	73	100.0	39	3	AAB17304	Aab17304 TPO-mimet
144	73	100.0	22	9	ADV44449	Adv44449 Anti-teta	217	73	100.0	39	3	AAB17305	Aab17305 TPO-mimet
145	73	100.0	22	9	ADV44443	Adv44443 Anti-teta	218	73	100.0	40	3	AAV96528	Aay96528 Thrombopo
146	73	100.0	22	9	ADV44444	Adv44444 Anti-teta	219	73	100.0	41	5	ABB73389	Abb73389 TPO-mimet
147	73	100.0	22	9	ADV44448	Adv44448 Anti-teta	220	73	100.0	41	5	ABB73388	Abb73388 TPO-mimet
148	73	100.0	22	9	ADV44442	Adv44442 Anti-teta	221	73	100.0	41	5	ABB73389	Abb73389 TPO-mimet
149	73	100.0	22	9	ADV44447	Adv44447 Anti-teta	222	73	100.0	42	3	AAV96530	Aay96530 Thrombopo
150	73	100.0	22	9	ADV44445	Adv44445 Anti-teta	223	73	100.0	42	3	AAV96530	Aay96530 Thrombopo
151	73	100.0	22	9	ADV44450	Adv44450 Anti-teta	224	73	100.0	42	3	AAB17308	Aab17308 Synthetic
152	73	100.0	22	9	ADV44446	Adv44446 Anti-teta	225	73	100.0	42	3	AAB17282	Aab17282 TPO-mimet
153	73	100.0	22	9	ABE12918	Aeb12918 TPO mimet	226	73	100.0	42	3	AAB17281	Aab17281 TPO-mimet
154	73	100.0	22	9	ABE12917	Aeb12917 TPO mimet	227	73	100.0	42	5	ABB73404	Abb73404 TMP-TMP g
155	73	100.0	22	9	ABE12919	Aeb12919 TPO mimet	228	73	100.0	60	3	AAB17311	Aab17311 Synthetic
156	73	100.0	22	9	ABE12986	Aeb12986 TT antibo	229	73	100.0	60	5	ABB73405	Abb73405 TMP-TMP g
157	73	100.0	22	9	ABE12920	Aeb12920 TPO mimet	230	73	100.0	122	9	ADV44474	Adv44474 Anti-teta
158	73	100.0	22	9	ABE12916	Aeb12916 TPO mimet	231	73	100.0	122	9	ADV44474	Adv44474 Anti-teta
159	73	100.0	22	9	ABE12914	Aeb12914 TPO mimet	232	73	100.0	128	8	ADQ16705	Adq16705 Modified
160	73	100.0	22	9	ABE12921	Aeb12921 TPO mimet	233	73	100.0	128	9	ADV44466	Adv44466 Anti-teta
161	73	100.0	22	9	ABE12907	Aeb12907 TT antibo	234	73	100.0	128	9	ADV44463	Adv44463 Anti-teta
162	73	100.0	22	9	ABE12915	Aeb12915 TPO mimet	235	73	100.0	128	9	ADV44467	Adv44467 Anti-teta
163	73	100.0	28	3	AAB17285	Aab17285 TPO-mimet	236	73	100.0	128	9	ADV44465	Adv44465 Anti-teta
164	73	100.0	28	5	ABP51682	Abp51682 TPO mimet	237	73	100.0	128	9	ADV44441	Adv44441 pAX116 va
165	73	100.0	28	7	ADJ73013	Adj73013 TPO mimet	238	73	100.0	128	9	ADV44464	Adv44464 Anti-teta
166	73	100.0	28	8	ADJ52648	Adj52648 CHI Delet	239	73	100.0	128	9	AEB12934	Aebi2934 Antibody
167	73	100.0	28	8	ADJ51609	Adj51609 CHI Delet	240	73	100.0	128	9	AEB12935	Aebi2935 Antibody
168	73	100.0	28	8	ADQ16636	Adq16636 Tetanus t	241	73	100.0	128	9	AEB12913	Aebi2913 Antibody
169	73	100.0	28	9	ADV44374	Adv44374 Modified	242	73	100.0	128	9	AEB12936	Aebi2936 Antibody
170	73	100.0	28	9	AEB12847	Aeb12847 Antibody	243	73	100.0	128	9	AEB12937	Aebi2937 Antibody

244	73	100.0	128	9	AB112938	Aeb12938	Antibody	317	67	91.8	15	8	ADM72496	Adm72496	TPO mimet
245	73	100.0	129	6	ABG71751	Abg71751	Antibody	318	67	91.8	15	8	ADM72494	Adm72494	TPO mimet
246	73	100.0	131	6	ABG71753	Abg71753	Antibody	319	66	90.4	13	4	AAU26008	Aau26008	Human thr
247	73	100.0	132	9	ADV44473	Adv44473	Anti-teta	320	65	89.0	12	2	AAW36787	Aaw36787	Thrombopo
248	73	100.0	133	6	ABG71752	Abg71752	Antibody	321	65	89.0	13	4	AAU26012	Aau26012	Human thr
249	73	100.0	135	6	ABG71749	Abg71749	Antibody	322	65	89.0	14	2	AAW36788	Aaw36788	Thrombopo
250	73	100.0	143	6	ABG71750	Abg71750	Antibody	323	65	89.0	14	4	AAU26013	Aau26013	Human thr
251	73	100.0	144	6	ABG71748	Abg71748	Antibody	324	65	89.0	15	2	AAW66717	Aaw66717	Peptide c
252	73	100.0	150	9	AB112945	Aeb12945	Antibody	325	64	87.7	12	2	AAW36781	Aaw36781	Thrombopo
253	73	100.0	225	8	ADQ16704	Adq16704	Modified	326	64	87.7	14	2	AAW36782	Aaw36782	Thrombopo
254	73	100.0	234	9	AB112912	Aeb12912	Antibody	327	64	87.7	14	4	AAU26009	Aau26009	Human thr
255	73	100.0	247	3	AB116958	Aab16958	FC-TMP pr	328	64	87.7	18	7	ADN59663	Adn59663	Thrombopo
256	73	100.0	247	3	AB116961	Aab16961	TMP-Fc pr	329	64	87.7	22	7	ADN59830	Adn59830	TMP pepti
257	73	100.0	247	3	AB116961	Aab16961	FC-TPO mi	330	64	87.7	25	7	ADN59708	Adn59708	Thrombopo
258	73	100.0	247	5	AB116961	Aab16961	TMP-Fc am	331	64	87.7	43	7	ADN59759	Adn59759	Peptide-v
259	73	100.0	249	5	ADV44440	Adv44440	PAX116 va	332	62	84.9	12	4	AAU26014	Aau26014	Human thr
260	73	100.0	249	5	ADV44440	Adv44440	FC-TMP-TM	333	61	83.6	14	4	AAU26037	Aau26037	Human thr
261	73	100.0	268	3	AB116959	Aab16959	FC-TMP-TM	334	61	83.6	14	8	ADM72526	Adm72526	TPO mimet
262	73	100.0	268	5	AB116959	Aab16959	FC-TMP-TM	335	61	83.6	14	8	ADM72499	Adm72499	TPO mimet
263	73	100.0	269	3	AAV96531	Aay96531	Human IgG	336	61	83.6	14	8	ADM72527	Adm72527	Peptide c
264	73	100.0	269	3	AAV96531	Aay96531	TMP-TMP-F	337	61	83.6	15	2	AAW66731	Aaw66731	Peptide c
265	73	100.0	269	3	AAV96531	Aay96531	TMP-TMP-F	338	61	83.6	15	4	AAU26024	Aau26024	Human thr
266	73	100.0	282	9	AB112930	Aeb12930	Antibody	339	61	83.6	15	4	AAU26038	Aau26038	Human thr
267	73	100.0	459	9	ADV44459	Adv44459	Anti-teta	340	61	83.6	15	8	ADM72498	Adm72498	TPO mimet
268	73	100.0	472	8	ABP51695	Abp51695	SG1.1-TPO	341	60	82.2	13	2	AAW36792	Aaw36792	Thrombopo
269	73	100.0	472	8	ADQ16647	Adq16647	Immunoglo	342	60	82.2	13	4	AAU26015	Aau26015	Human thr
270	73	100.0	472	9	ADV44385	Adv44385	SG1.1 hea	343	60	82.2	14	8	ADM72475	Adm72475	TPO mimet
271	73	100.0	472	9	AB112958	Aeb12958	Antibody	344	60	82.2	14	9	ADM72477	Adm72477	Thrombopo
272	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	345	60	82.2	15	9	ADY64335	Ady64335	Thrombopo
273	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	346	60	82.2	15	9	ADY64335	Ady64335	Thrombopo
274	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	347	60	82.2	16	2	AAW66711	Aaw66711	Peptide c
275	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	348	60	82.2	16	8	ADM72476	Adm72476	TPO mimet
276	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	349	60	82.2	19	2	AAW09491	Aaw09491	Thrombopo
277	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	350	60	82.2	19	2	AAW09493	Aaw09493	Thrombopo
278	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	351	60	82.2	19	2	AAW36644	Aaw36644	Thrombopo
279	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	352	60	82.2	19	2	AAW36644	Aaw36644	Thrombopo
280	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	353	60	82.2	19	2	AAW36644	Aaw36644	Thrombopo
281	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	354	60	82.2	19	4	AAU25861	Aau25861	Human thr
282	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	355	60	82.2	19	4	AAU25863	Aau25863	Human thr
283	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	356	60	82.2	19	4	AAU25998	Aau25998	Human thr
284	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	357	60	82.2	28	8	ADJ52647	Adj52647	CH1 delet
285	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	358	60	82.2	29	3	AA116972	Aab16972	TPO-mimet
286	73	100.0	95.9	14	AA116968	Aab16968	TPO-mimet	359	60	82.2	29	5	ABB72858	Abb72858	TPO mimet
287	69	94.5	13	2	AAW36779	Aaw36779	Thrombopo	360	60	82.2	29	7	ADJ73012	Adj73012	TPO mimet
288	69	94.5	13	7	ADJ73003	Adj73003	TPO mimet	361	60	82.2	29	8	ADJ73008	Adj73008	TPO mimet
289	69	94.5	13	8	ADJ52638	Adj52638	CH1 delet	362	60	82.2	29	8	ADJ52643	Adj52643	CH1 delet
290	69	94.5	13	8	ADJ52638	Adj52638	CH1 delet	363	60	82.2	29	8	ADJ51604	Adj51604	CH1 delet
291	69	94.5	14	4	AAU26006	Aau26006	Human thr	364	60	82.2	29	8	ADJ51608	Adj51608	CH1 delet
292	69	94.5	14	4	AAU26010	Aau26010	Human thr	365	60	82.2	36	3	AAV96521	Aay96521	Cyclic or
293	69	94.5	14	8	ADM72503	Adm72503	TPO mimet	366	60	82.2	36	3	AAV96521	Aay96521	Cyclic or
294	69	94.5	14	8	ADM72487	Adm72487	TPO mimet	367	60	82.2	36	3	AAV96521	Aay96521	Cyclic or
295	69	94.5	14	9	ADU70206	Adu70206	Thrombopo	368	59	80.8	14	8	ADM72509	Adm72509	TPO-mimet
296	69	94.5	14	9	ADU75979	Adu75979	Thrombopo	369	59	80.8	15	2	AAW66723	Aaw66723	Peptide c
297	69	94.5	15	2	AAW36784	Aaw36784	Thrombopo	370	59	80.8	15	4	AAU26028	Aau26028	Human thr
298	69	94.5	15	2	AAW36780	Aaw36780	Thrombopo	371	59	80.8	15	8	ADM72508	Adm72508	TPO mimet
299	69	94.5	15	2	AAW66714	Aaw66714	Peptide c	372	59	80.8	18	2	AAW09460	Aaw09460	Thrombopo
300	69	94.5	15	2	AAW66721	Aaw66721	Peptide c	373	59	80.8	18	2	AAW09498	Aaw09498	Thrombopo
301	69	94.5	15	4	AAU26026	Aau26026	Human thr	374	59	80.8	18	2	AAW36649	Aaw36649	Thrombopo
302	69	94.5	15	4	AAU26011	Aau26011	Human thr	375	59	80.8	18	2	AAW33027	Aaw33027	Thrombopo
303	69	94.5	15	4	AAU26020	Aau26020	Human thr	376	59	80.8	18	2	AAW36652	Aaw36652	Thrombopo
304	69	94.5	15	4	AAU26007	Aau26007	Human thr	377	59	80.8	18	3	AAW17026	Aab17026	TPO-mimet
305	69	94.5	15	8	ADM72502	Adm72502	TPO mimet	378	59	80.8	18	4	AAU25868	Aau25868	Human thr
306	69	94.5	15	8	ADM72492	Adm72492	TPO mimet	379	59	80.8	18	4	AAU25824	Aau25824	Human thr
307	69	94.5	15	8	ADM72490	Adm72490	TPO mimet	380	59	80.8	18	4	AAU25871	Aau25871	Human thr
308	69	94.5	15	8	ADM72486	Adm72486	TPO mimet	381	59	80.8	18	5	ABB72312	Abb72312	TPO mimet
309	69	94.5	15	8	ADM72491	Adm72491	TPO mimet	382	59	80.8	18	8	ADJ52699	Adj52699	CH1 delet
310	69	94.5	15	8	ADM72493	Adm72493	TPO mimet	383	59	80.8	18	8	ADJ51660	Adj51660	CH1 delet
311	69	94.5	16	4	AAU26021	Aau26021	Human thr	384	59	80.8	36	3	AAV96522	Aay96522	Linear th
312	67	91.8	14	8	ADM72495	Adm72495	TPO mimet	385	59	80.8	36	3	AAV96522	Aay96522	Linear th
313	67	91.8	14	8	ADM72497	Adm72497	TPO mimet	386	59	80.8	36	3	AAV96522	Aay96522	Linear th
314	67	91.8	15	2	AAW66719	Aaw66719	Peptide c	387	58	79.5	13	2	AAW36783	Aaw36783	Thrombopo
315	67	91.8	15	4	AAU26022	Aau26022	Human thr	388	58	79.5	13	4	AAU26035	Aau26035	Human thr
316	67	91.8	15	4	AAU26023	Aau26023	Human thr	389	58	79.5	13	8	ADM72524	Adm72524	TPO mimet

390	58	79.5	14	2	AAW36773	Thrombopo	463	53	72.6	19	2	AAW09492	Thrombopo
391	57	78.1	13	4	AAU26018	Human thr	464	53	72.6	19	2	AAW36651	Thrombopo
392	57	78.1	13	8	ADM72488	TPO minet	465	53	72.6	19	2	AAW33024	Thrombopo
393	57	78.1	18	2	AAW09499	Thrombopo	466	53	72.6	19	2	AAW36643	Thrombopo
394	57	78.1	18	2	AAW09459	Thrombopo	467	53	72.6	19	3	AAAB17021	TPO-minet
395	57	78.1	18	2	AAW36650	Thrombopo	468	53	72.6	19	4	AAU25862	Human thr
396	57	78.1	18	2	AAW33026	Thrombopo	469	53	72.6	19	4	AAU25870	Human thr
397	57	78.1	18	3	AAAB17024	TPO-minet	470	53	72.6	19	4	AAU25821	Human thr
398	57	78.1	18	4	AAU25869	Human thr	471	53	72.6	19	5	ABB72907	TPO minet
399	57	78.1	18	4	AAU25823	Human thr	472	53	72.6	19	7	ADJ73059	TPO minet
400	57	78.1	18	5	ABB72910	TPO minet	473	53	72.6	19	8	ADJ52694	CH1 delet
401	57	78.1	18	7	ADJ73062	TPO minet	474	53	72.6	19	8	ADJ51655	CH1 delet
402	57	78.1	18	8	ADJ52697	CH1 delet	475	53	72.6	25	7	ADN59742	Thrombopo
403	57	78.1	18	8	ADJ52697	CH1 delet	476	53	72.6	25	7	ADN59730	Thrombopo
404	56	76.7	10	3	AAAB17006	TPO-minet	477	53	72.6	25	7	ADN59740	Thrombopo
405	56	76.7	10	5	ABB72892	TPO minet	478	53	72.6	43	7	ADN59752	Peptide-v
406	56	76.7	10	7	ADJ73043	TPO minet	479	52	71.2	12	3	AAAB17309	Synthetic
407	56	76.7	10	8	ADJ52678	CH1 delet	480	52	71.2	18	7	ADN59653	Thrombopo
408	56	76.7	10	8	ADJ51639	CH1 delet	481	52	71.2	18	7	ADN59670	Thrombopo
409	56	76.7	10	8	ADJ52678	CH1 delet	482	52	71.2	18	7	ADN59667	Thrombopo
410	56	76.7	13	8	ADM72489	TPO minet	483	52	71.2	22	7	ADN59820	TMP pepti
411	56	76.7	13	8	AAW66715	Peptide c	484	52	71.2	22	7	ADN59837	TMP pepti
412	56	76.7	14	2	AAW66730	Peptide c	485	52	71.2	22	7	ADN59837	TMP pepti
413	56	76.7	14	2	AAU26019	Human thr	486	52	71.2	25	7	ADN59689	Thrombopo
414	56	76.7	14	4	AAU26036	Human thr	487	52	71.2	25	7	ADN59722	Thrombopo
415	56	76.7	14	8	ADM72507	TPO minet	488	52	71.2	25	7	ADN59716	Thrombopo
416	56	76.7	14	8	ADM72505	TPO minet	489	52	71.2	42	7	ADN59818	Peptide-
417	56	76.7	14	8	ADM72501	TPO minet	490	51	69.9	14	8	ADM72511	TPO minet
418	56	76.7	14	8	ADU70205	Thrombopo	491	51	69.9	14	8	ADM72519	TPO minet
419	56	76.7	14	9	ADU75978	Thrombopo	492	51	69.9	15	2	AAW66728	Peptide c
420	56	76.7	15	2	AAW66722	Peptide c	493	51	69.9	15	4	AAU26033	Human thr
421	56	76.7	15	2	AAW66720	Peptide c	494	51	69.9	15	4	AAU26029	Human thr
422	56	76.7	15	4	AAU25833	Human thr	495	51	69.9	15	8	ADM72510	TPO minet
423	56	76.7	15	4	AAU26027	Human thr	496	51	69.9	15	8	ADM72518	TPO minet
424	56	76.7	15	4	AAU26025	Human thr	497	51	69.9	18	7	ADN59654	Thrombopo
425	56	76.7	15	8	ADM72506	TPO minet	498	51	69.9	18	7	ADN59678	Thrombopo
426	56	76.7	15	8	ADM72500	TPO minet	499	51	69.9	18	7	ADN59669	Thrombopo
427	56	76.7	15	8	ADM72504	TPO minet	500	51	69.9	22	7	ADN59836	TMP pepti
428	56	76.7	18	7	ADN59664	Thrombopo	501	51	69.9	22	7	ADN59821	TMP pepti
429	56	76.7	18	7	ADN59666	Thrombopo	502	51	69.9	25	7	ADN59736	Thrombopo
430	56	76.7	19	2	AAW09494	Thrombopo	503	51	69.9	25	7	ADN59691	Thrombopo
431	56	76.7	19	2	AAW09461	Thrombopo	504	51	69.9	25	7	ADN59720	Thrombopo
432	56	76.7	19	2	AAW33028	Thrombopo	505	50	68.5	10	2	AAW09469	Thrombopo
433	56	76.7	19	2	AAW36645	Thrombopo	506	50	68.5	10	2	AAW36621	Thrombopo
434	56	76.7	19	3	AAAB17022	TPO-minet	507	50	68.5	10	2	AAW36772	Thrombopo
435	56	76.7	19	4	AAU25864	Human thr	508	50	68.5	10	3	AAAB17005	TPO-minet
436	56	76.7	19	4	AAU25825	Human thr	509	50	68.5	10	4	AAU25839	Human thr
437	56	76.7	19	5	ABB72908	TPO minet	510	50	68.5	10	4	AAU25840	Human thr
438	56	76.7	19	7	ADJ73060	TPO minet	511	50	68.5	10	5	ABB72891	TPO minet
439	56	76.7	19	8	ADJ52695	CH1 delet	512	50	68.5	10	7	ADJ73042	TPO minet
440	56	76.7	19	8	ADJ51656	CH1 delet	513	50	68.5	10	8	ADJ52677	CH1 delet
441	56	76.7	22	7	ADN59831	TMP pepti	514	50	68.5	10	8	ADJ51638	CH1 delet
442	56	76.7	22	7	ADN59827	TMP pepti	515	50	68.5	15	2	AAW66725	Peptide c
443	56	76.7	25	7	ADN59710	Thrombopo	516	50	68.5	18	7	ADN59657	Thrombopo
444	56	76.7	25	7	ADN59702	Thrombopo	517	50	68.5	22	7	ADN59824	TMP pepti
445	56	76.7	30	9	ADY64336	Thrombopo	518	50	68.5	23	7	ADN59775	Peptide-v
446	56	76.7	42	7	ADN59751	Peptide-v	519	50	68.5	23	7	ADN59793	Peptide-v
447	54	74.0	18	7	ADN59815	Thrombopo	520	50	68.5	25	7	ADN59696	Thrombopo
448	54	74.0	18	7	ADN59668	Thrombopo	521	50	68.5	36	7	ADN59763	Peptide-v
449	54	74.0	18	7	ADN59666	Thrombopo	522	50	68.5	41	7	ADN59769	Peptide-v
450	54	74.0	22	7	ADN59835	TMP pepti	523	50	68.5	46	7	ADN59781	Peptide-v
451	54	74.0	22	7	ADN59833	TMP pepti	524	50	68.5	46	7	ADN59787	Peptide-v
452	54	74.0	25	7	ADN59718	Thrombopo	525	50	68.5	302	6	ABU33828	Protein a
453	54	74.0	25	7	ADN59714	Thrombopo	526	49	67.1	10	2	AAW09472	Thrombopo
454	53	72.6	15	2	AAW66724	Peptide c	527	49	67.1	10	2	AAW36623	Thrombopo
455	53	72.6	15	4	AAU26016	Human thr	528	49	67.1	10	4	AAU25842	Human thr
456	53	72.6	15	8	ADM72480	TPO minet	529	49	67.1	11	2	AAW36791	Thrombopo
457	53	72.6	15	8	ADM72481	TPO minet	530	49	67.1	11	3	AAAB17015	TPO-minet
458	53	72.6	16	4	ADM726017	Human thr	531	49	67.1	13	5	ABB72901	TPO minet
459	53	72.6	18	7	ADN59681	Thrombopo	532	49	67.1	13	7	ADJ73054	TPO minet
460	53	72.6	18	7	ADN59675	Thrombopo	533	49	67.1	13	7	ADJ73052	TPO minet
461	53	72.6	18	7	ADN59680	Thrombopo	534	49	67.1	13	7	ADJ73056	TPO minet
462	53	72.6	19	2	AAW09457	Thrombopo	535	49	67.1	13	7	ADJ73053	TPO minet

682	45	61.6	14	5	ABB72900	Abb72900 TPO mimet	755	42	57.5	15	8	ADM72516	Adm72516 TPO mimet
683	45	61.6	14	7	ADJ73051	Adj73051 TPO mimet	756	42	57.5	18	2	AAW09592	Aaw09592 Thrombopo
684	45	61.6	14	8	ADJ52686	Adj52686 CH1 delet	757	42	57.5	18	2	AAW36743	AAW36743 Thrombopo
685	45	61.6	14	8	ADJ51647	Adj51647 CH1 delet	758	42	57.5	18	4	AAU25962	Aau25962 Human thr
686	45	61.6	14	8	ADM72515	Adm72515 TPO mimet	759	42	57.5	100	5	AAU25962	AAU25962 Human thr
687	45	61.6	14	8	ADM72521	Adm72521 TPO mimet	760	42	57.5	110	4	AAW85655	AAW85655 Human imm
688	45	61.6	15	2	AAW66729	Aaw66729 Peptide c	761	42	57.5	126	7	ADM03803	Adm03803 Human pro
689	45	61.6	15	4	AAU26031	Aau26031 Human thr	762	42	57.5	429	5	ABP47758	Abp47758 Protein #
690	45	61.6	15	4	AAU26034	Aau26034 Human thr	763	42	57.5	450	3	AAW50584	AAW50584 Arabidops
691	45	61.6	15	8	ADM72520	Adm72520 TPO mimet	764	42	57.5	475	3	AAW50584	AAW50584 Arabidops
692	45	61.6	15	8	ADM72514	Adm72514 TPO mimet	765	42	57.5	496	9	ABM91663	ABM91663 M. xanthu
693	45	61.6	18	2	AAW09456	Aaw09456 Thrombopo	766	42	57.5	499	3	AAW50582	AAW50582 Arabidops
694	45	61.6	18	2	AAW09487	Aaw09487 Thrombopo	767	42	57.5	512	4	ABG25367	ABG25367 Novel hum
695	45	61.6	18	2	AAW33023	Aaw33023 Thrombopo	768	42	57.5	519	7	ADM05082	Adm05082 Human pro
696	45	61.6	18	2	AAW36638	Aaw36638 Thrombopo	769	41	56.2	18	2	AAW09490	AAW09490 Human pro
697	45	61.6	18	3	AAAB17020	Aaab17020 TPO-mimet	770	41	56.2	18	2	AAW36641	AAW36641 Thrombopo
698	45	61.6	18	4	AAU25857	Aau25857 Human thr	771	41	56.2	18	4	AAU25860	Aau25860 Human thr
699	45	61.6	18	4	AAU25820	Aau25820 Human thr	772	41	56.2	18	7	ADN59676	Adn59676 Thrombopo
700	45	61.6	18	5	ABB72906	Abb72906 TPO mimet	773	41	56.2	18	7	ADN59662	Adn59662 Thrombopo
701	45	61.6	18	7	ADJ73058	Adj73058 TPO mimet	774	41	56.2	22	7	ADN59829	Adn59829 TMP pepti
702	45	61.6	18	8	ADJ52693	Adj52693 CH1 delet	775	41	56.2	25	7	ADN59706	Adn59706 Thrombopo
703	45	61.6	18	8	ADJ51654	Adj51654 CH1 delet	776	41	56.2	25	7	ADN59732	Adn59732 Thrombopo
704	45	61.6	19	2	AAW09458	Aaw09458 Thrombopo	777	41	56.2	49	9	ABM94741	ABM94741 M. xanthu
705	45	61.6	19	2	AAW33025	Aaw33025 Human thr	778	41	56.2	137	8	ABO84792	ABO84792 Murine ca
706	45	61.6	19	4	AAU25822	Aau25822 Human thr	779	41	56.2	155	3	AAW12015	AAW12015 Arabidops
707	44.5	61.0	316	9	ABM93583	ABm93583 M. xanthu	780	41	56.2	211	3	ABA43611	ABa43611 Human can
708	44	60.3	13	8	ADM72530	Adm72530 TPO mimet	781	41	56.2	211	4	ABG22919	ABg22919 Novel hum
709	44	60.3	13	4	AAU26041	Aau26041 Human thr	782	41	56.2	220	4	ABG18864	ABg18864 Novel hum
710	44	60.3	13	8	ADM72529	Adm72529 TPO mimet	783	41	56.2	291	3	AAW25112	AAW25112 Arabidops
711	44	60.3	14	2	AAW66732	Aaw66732 Peptide c	784	41	56.2	294	3	AAW25111	AAW25111 Arabidops
712	44	60.3	14	4	AAU26040	Aau26040 Human thr	785	41	56.2	294	3	AAW55002	AAW55002 Arabidops
713	44	60.3	15	2	AAW66727	Aaw66727 Peptide c	786	41	56.2	294	3	AAW17087	AAW17087 Arabidops
714	44	60.3	18	2	AAW09497	Aaw09497 Thrombopo	787	41	56.2	306	5	ABG51631	ABg51631 Putine/py
715	44	60.3	18	2	AAW36653	Aaw36653 Thrombopo	788	41	56.2	306	6	ABU35738	ABu35738 Protein e
716	44	60.3	18	2	AAW36648	Aaw36648 Thrombopo	789	41	56.2	308	4	AAW91969	AAW91969 C glutami
717	44	60.3	18	3	AAW17025	Aaw17025 TPO-mimet	790	41	56.2	311	7	ABM89301	ABm89301 Rice abio
718	44	60.3	18	4	AAU25872	Aau25872 Human thr	791	41	56.2	312	8	ADT56484	ADt56484 Plant pol
719	44	60.3	18	4	AAU25867	Aau25867 Human thr	792	41	56.2	318	8	ADY22882	ADy22882 Plant ful
720	44	60.3	18	5	ABB72911	Abb72911 TPO mimet	793	41	56.2	326	1	AAW82999	AAW82999 Tobacco G
721	44	60.3	18	7	ADJ73063	Adj73063 TPO mimet	794	41	56.2	338	3	AAW17086	AAW17086 Arabidops
722	44	60.3	18	7	ADN59674	Adn59674 Thrombopo	795	41	56.2	338	3	AAW25110	AAW25110 Arabidops
723	44	60.3	18	8	ADJ52698	Adj52698 CH1 delet	796	41	56.2	364	3	AAW17085	AAW17085 Arabidops
724	44	60.3	18	8	ADJ51659	Adj51659 CH1 delet	797	41	56.2	876	9	ABM92266	ABm92266 M. xanthu
725	44	60.3	22	7	ADN59841	Adn59841 TMP pepti	798	40	54.8	14	2	AAW09484	AAW09484 Thrombopo
726	44	60.3	23	7	ADN59797	Adn59797 Peptide-v	799	40	54.8	14	2	AAW36766	AAW36766 Thrombopo
727	44	60.3	23	7	ADN59779	Adn59779 Peptide-v	800	40	54.8	14	2	AAW36635	AAW36635 Thrombopo
728	44	60.3	25	7	ADN59728	Adn59728 Thrombopo	801	40	54.8	14	3	AAW17009	AAW17009 TPO-mimet
729	44	60.3	35	7	ADN59754	Adn59754 Peptide-v	802	40	54.8	14	4	AAU25854	Aau25854 Human thr
730	44	60.3	36	7	ADN59767	Adn59767 Peptide-v	803	40	54.8	14	5	ABB72895	ABb72895 TPO mimet
731	44	60.3	41	7	ADN59773	Adn59773 Peptide-v	804	40	54.8	14	7	ADJ73046	ADj73046 TPO mimet
732	44	60.3	46	7	ADN59785	Adn59785 Peptide-v	805	40	54.8	14	8	ADJ52681	ADj52681 CH1 delet
733	44	60.3	46	7	ADN59791	Adn59791 Peptide-v	806	40	54.8	14	8	ADJ51642	ADj51642 CH1 delet
734	44	60.3	75	7	ADN59758	Adn59758 Peptide-v	807	40	54.8	19	2	AAW09495	AAW09495 Thrombopo
735	44	60.3	87	4	AAU63256	Aau63256 Propionib	808	40	54.8	19	2	AAW36646	AAW36646 Thrombopo
736	44	60.3	87	6	ABM59775	ABm59775 Propionib	809	40	54.8	19	3	AAW17023	AAW17023 TPO-mimet
737	44	60.3	135	4	ABG229134	ABg229134 Novel hum	810	40	54.8	19	4	AAU25865	Aau25865 Human thr
738	44	60.3	200	8	ADN23470	Adn23470 Bacterial	811	40	54.8	19	5	ABB72909	ABb72909 TPO mimet
739	44	60.3	536	6	ABU15399	ABu15399 Protein e	812	40	54.8	19	7	ADJ73061	ADj73061 TPO mimet
740	44	60.3	536	8	ADK13781	ADk13781 E. coli i	813	40	54.8	19	8	ADJ52696	ADj52696 CH1 delet
741	44	60.3	536	8	ADK18093	ADk18093 Bacterial	814	40	54.8	19	8	ADJ51657	ADj51657 CH1 delet
742	43	58.9	11	2	AAW35425	Aaw35425 Thrombopo	815	40	54.8	28	7	AAE38642	AAe38642 Human col
743	43	58.9	11	4	AAU26001	Aau26001 Human thr	816	40	54.8	153	9	ABM92349	ABm92349 M. xanthu
744	43	58.9	18	2	AAW09488	Aaw09488 Thrombopo	817	40	54.8	207	6	ABU34877	ABu34877 Protein e
745	43	58.9	18	2	AAW36639	Aaw36639 Thrombopo	818	40	54.8	239	8	ADM80079	ADM80079 Spiramy
746	43	58.9	18	4	AAU25858	Aau25858 Human thr	819	40	54.8	239	8	ADN97595	ADn97595 S ambofac
747	43	58.9	18	7	ADN59679	Adn59679 Thrombopo	820	40	54.8	261	7	ADH86920	ADh86920 Enterococ
748	43	58.9	18	7	ADN59677	Adn59677 Thrombopo	821	40	54.8	271	8	ADS42097	ADs42097 Bacterial
749	43	58.9	25	7	ADN59734	Adn59734 Thrombopo	822	40	54.8	311	4	AAW41798	AAW41798 Human pol
750	43	58.9	25	7	ADN59738	Adn59738 Thrombopo	823	40	54.8	329	4	ABG23127	ABg23127 Novel hum
751	43	58.9	40	7	ADN59760	Adn59760 Peptide-v	824	40	54.8	355	6	ABR58696	ABr58696 Human can
752	43	58.9	285	6	ABU48553	ABu48553 Protein e	825	40	54.8	355	7	ADF13714	ADF13714 Tumor-Aas
753	42	57.5	14	8	ADM72517	Adm72517 TPO mimet	826	40	54.8	355	8	ADS00100	ADs00100 Human ocu
754	42	57.5	15	4	AAU26032	Aau26032 Human thr	827	40	54.8	355	8	ADS00098	ADs00098 Human ocu

828	40	54.8	355	9	ADW14773	Adw14773 Tumor-agg	901	38	52.1	10	2	AAW36624	AAW36624 Thrombopo
829	40	54.8	410	6	ABU22259	Abu22259 Protein e	902	38	52.1	10	4	AAU25843	AAU25843 Human thr
830	40	54.8	410	6	ABU22026	Abu22026 Protein e	903	38	52.1	89	4	AAW83690	AAW83690 Human ova
831	40	54.8	486	6	ABP77210	Abp77210 N. gonorr	904	38	52.1	126	2	AAW76522	AAW76522 Human pro
832	40	54.8	492	7	ABO82111	AbO82111 Pseudomon	905	38	52.1	132	7	ADM04387	ADM04387 Arabidops
833	40	54.8	513	2	AAW38788	Aay38788 Neisseria	906	38	52.1	161	3	AAW28444	AAW28444 Arabidops
834	40	54.8	513	2	AAW38786	Aay38786 Neisseria	907	38	52.1	162	8	ADW94945	ADW94945 Plant ful
835	40	54.8	513	2	AAW38787	Aay38787 Neisseria	908	38	52.1	174	8	ADW91393	ADW91393 Plant ful
836	40	54.8	513	2	AAW38788	Aay38788 Neisseria	909	38	52.1	206	3	AAW28443	AAW28443 Arabidops
837	40	54.8	513	9	ABE49392	Aeb49392 N. mening	910	38	52.1	212	7	ABO68391	ABO68391 Pseudomon
838	40	54.8	513	9	ABE49394	Aeb49394 N. gonorr	911	38	52.1	232	8	ADW95040	ADW95040 Plant ful
839	40	54.8	513	9	ABE49395	Aeb49395 N. gonorr	912	38	52.1	240	8	ADW10154	ADW10154 Plant ful
840	40	54.8	513	9	ABE49397	Aeb49397 N. gonorr	913	38	52.1	243	7	ADW71135	ADW71135 Human hea
841	40	54.8	514	7	ABO82206	AbO82206 Pseudomon	914	38	52.1	244	6	ABU22180	ABU22180 Protein e
842	40	54.8	549	2	AAW40113	Aaw40113 Human alp	915	38	52.1	270	8	ADW92531	ADW92531 Plant ful
843	40	54.8	603	8	ADW25043	Adw25043 Bacterial	916	38	52.1	310	9	ABM96497	ABM96497 M. xanthu
844	40	54.8	656	8	ADW71949	Adw71949 Human PWM	917	38	52.1	312	6	ABM67747	ABM67747 Phototrab
845	40	54.8	686	8	ABM84061	Abm84061 Human dia	918	38	52.1	313	7	ADW04965	ADW04965 Bacterial
846	40	54.8	699	9	ADW85513	Adw85513 Human tra	919	38	52.1	321	8	ADW91347	ADW91347 Plant ful
847	40	54.8	721	9	ADW26477	Adw26477 Human ery	920	38	52.1	327	8	ADW91347	ADW91347 Plant ful
848	40	54.8	721	8	ABM82875	Abm82875 Human dia	921	38	52.1	333	8	ADW68901	ADW68901 Plant ful
849	40	54.8	994	8	ABM82874	Abm82874 Human dia	922	38	52.1	334	7	ABO65391	ABO65391 Klebsiell
850	40	54.8	1061	8	AAW40012	Aaw40012 Human poi	923	38	52.1	334	8	ADW95928	ADW95928 Plant ful
851	40	54.8	1078	5	AAO17361	Aao17361 Human alp	924	38	52.1	338	8	ADT57351	ADT57351 Plant pol
852	40	54.8	1712	8	ADG10473	Adg10473 Human the	925	38	52.1	338	8	ADW09292	ADW09292 Plant ful
853	40	54.8	1712	8	ADW06526	Adw06526 Novel bro	926	38	52.1	339	8	ADW14943	ADW14943 Pseudomon
854	40	54.8	1712	9	ADW59885	Adw59885 Human col	927	38	52.1	343	8	ADW96259	ADW96259 Plant ful
855	39.5	54.1	571	4	ABG29052	Abg29052 Novel hum	928	38	52.1	344	8	ADY10815	ADY10815 Plant ful
856	39.5	54.1	821	4	ABG29013	Abg29013 Novel hum	929	38	52.1	344	8	ADW96259	ADW96259 Plant ful
857	39.5	54.1	821	4	ABG29872	Abg29872 Novel hum	930	38	52.1	344	8	ADW95200	ADW95200 Plant ful
858	39	53.4	14	3	AAW17010	Aaw17010 TPO-mimet	931	38	52.1	344	8	ADW96519	ADW96519 Plant ful
859	39	53.4	14	3	ABW72896	Abw72896 TPO mimet	932	38	52.1	344	8	ADY08650	ADY08650 Plant ful
860	39	53.4	14	7	ADW73047	Adw73047 CHI delet	933	38	52.1	344	8	ADY13023	ADY13023 Plant ful
861	39	53.4	14	8	ADW52682	Adw52682 CHI delet	934	38	52.1	344	8	ADW09844	ADW09844 Plant ful
862	39	53.4	14	8	ADW51643	Adw51643 CHI delet	935	38	52.1	344	8	ADY12363	ADY12363 Plant ful
863	39	53.4	18	2	AAW09489	Aaw09489 Thrombopo	936	38	52.1	344	8	ADW98992	ADW98992 Plant ful
864	39	53.4	18	2	AAW36640	Aaw36640 Thrombopo	937	38	52.1	344	8	ADY11386	ADY11386 Plant ful
865	39	53.4	18	2	AAW25859	Aau25859 Human thr	938	38	52.1	345	8	ADY12305	ADY12305 Plant ful
866	39	53.4	95	5	ABP06014	Abp06014 Human ORF	939	38	52.1	349	8	ADW76194	ADW76194 Plant ful
867	39	53.4	95	5	AAU09166	Aau09166 Human cyt	940	38	52.1	350	2	AAW60244	AAW60244 Amino aci
868	39	53.4	100	5	ABP47800	Abp47800 HIV-2prot	941	38	52.1	350	4	AAU69755	AAU69755 Escherich
869	39	53.4	128	4	AAU57127	Aau57127 Propionib	942	38	52.1	351	8	ADW95138	ADW95138 Plant ful
870	39	53.4	128	6	ABW53646	Abw53646 Propionib	943	38	52.1	351	8	ADW06579	ADW06579 Plant ful
871	39	53.4	168	4	ABW65448	Abw65448 Drosophil	944	38	52.1	355	8	ADW06836	ADW06836 Plant ful
872	39	53.4	201	8	ADN25936	Adn25936 Bacterial	945	38	52.1	355	8	ADW93406	ADW93406 Plant ful
873	39	53.4	230	2	AAW60767	Aaw60767 HIV-2 ROD	946	38	52.1	356	8	ADY08052	ADY08052 Plant ful
874	39	53.4	267	5	AAU91162	Aau91162 Bordella	947	38	52.1	357	8	ADW66849	ADW66849 Plant ful
875	39	53.4	318	9	ABM91698	Abm91698 M. xanthu	948	38	52.1	357	8	ADW67086	ADW67086 Plant ful
876	39	53.4	321	8	ADW65923	Adw65923 Plant ful	949	38	52.1	357	8	ADW68311	ADW68311 Plant ful
877	39	53.4	324	8	ADW22836	Adw22836 Bacterial	950	38	52.1	361	8	ADW67626	ADW67626 Plant ful
878	39	53.4	397	6	ABR82135	AbR82135 Bacterial	951	38	52.1	361	8	ADW94971	ADW94971 Plant ful
879	39	53.4	412	7	ABM87926	Abm87926 Rice abio	952	38	52.1	362	8	ADW93509	ADW93509 Plant ful
880	39	53.4	428	5	ABP47757	Abp47757 Protein #	953	38	52.1	362	8	ADW93520	ADW93520 Plant ful
881	39	53.4	476	8	ADT58335	Adt58335 Plant pol	954	38	52.1	363	8	ADW11959	ADW11959 Plant ful
882	39	53.4	497	8	ADT57737	Adt57737 Plant pol	955	38	52.1	364	8	ADW13598	ADW13598 Plant ful
883	39	53.4	517	8	ADW23040	Adw23040 Plant ful	956	38	52.1	364	8	ADY06580	ADY06580 Plant ful
884	39	53.4	526	8	ADY07792	Ady07792 Plant ful	957	38	52.1	364	8	ADY22675	ADY22675 Plant ful
885	39	53.4	619	7	ADW87640	Adw87640 Enterococ	958	38	52.1	364	8	ADW80999	ADW80999 HIV prote
886	39	53.4	706	6	AAW22341	Aaw22341 Protein e	959	38	52.1	373	4	ADW10783	ADW10783 Plant ful
887	39	53.4	764	1	AAW81784	Aaw81784 P. putida	960	38	52.1	373	8	ADW10783	ADW10783 Plant ful
888	39	53.4	1014	1	AAW81773	Aap81773 Sequence	961	38	52.1	389	6	ABO76974	ABO76974 Pseudomon
889	39	53.4	1027	1	AAW81773	Aap81773 Sequence	962	38	52.1	489	7	ABO76974	ABO76974 Pseudomon
890	39	53.4	1035	2	AAW81773	Aap81773 Sequence	963	38	52.1	489	7	ABO76974	ABO76974 Pseudomon
891	39	53.4	1036	2	AAW81773	Aap81773 Sequence	964	38	52.1	496	4	ABG99875	ABG99875 S. cinnam
892	39	53.4	1036	3	AAW81773	Aap81773 Sequence	965	38	52.1	496	4	ABG99875	ABG99875 S. cinnam
893	39	53.4	1036	3	AAW81773	Aap81773 Sequence	966	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
894	39	53.4	1036	3	AAW81773	Aap81773 Sequence	967	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
895	39	53.4	1256	7	ADW23699	Adw23699 Murine ne	968	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
896	39	53.4	1275	6	ABJ25467	Abj25467 Aspergill	969	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
897	39	53.4	1275	6	ABJ25467	Abj25467 Aspergill	970	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
898	39	53.4	1435	8	ADN27223	Adn27223 Bacterial	971	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
899	39	53.4	1437	8	ADN27223	Adn27223 Bacterial	972	38	52.1	572	7	ABO80566	ABO80566 Pseudomon
900	38	52.1	10	2	AAW09473	Aaw09473 Thrombopo	973	38	52.1	599	9	ADW40028	ADW40028 HIV repli p

974 38 52.1 1252 7 ADE97411 Norway ra
975 38 52.1 1256 4 AB84865 Murine ne
976 38 52.1 3781 4 AB871980 Drosophil
977 37.5 51.4 131 7 ABO67862 Pseudomon
978 37.5 51.4 169 7 ABO76480 Pseudomon
979 37.5 51.4 719 7 ABO74914 Pseudomon
980 37 50.7 10 2 AAW35427 Thrombopo
981 37 50.7 10 2 AAW26002 Human thr
982 37 50.7 14 2 AAW09479 Thrombopo
983 37 50.7 14 2 AAW09481 Thrombopo
984 37 50.7 14 2 AAW36630 Thrombopo
985 37 50.7 14 2 AAW36632 Thrombopo
986 37 50.7 14 2 AAW36631 Thrombopo
987 37 50.7 14 3 AAB17013 TPO-mimet
988 37 50.7 14 4 AAU25849 Human thr
989 37 50.7 14 4 AAU25850 Human thr
990 37 50.7 14 4 AAU25851 Human thr
991 37 50.7 14 5 ABB72899 TPO mimet
992 37 50.7 14 7 ADJ73050 TPO mimet
993 37 50.7 14 8 ADJ52685 CH1 delet
994 37 50.7 14 8 ADJ51646 CH1 delet
995 37 50.7 54 4 AAU54078 Propionib
996 37 50.7 54 6 ABM50597 Propionib
997 37 50.7 100 5 ABP47799 SIV prote
998 37 50.7 108 4 AAM84222 Human imm
999 37 50.7 108 4 AAU57517 Propionib
1000 37 50.7 108 6 ABM54036 Propionib

ALIGNMENTS

RESULT 1
ID AAW09463 standard; protein; 14 AA.

AC AAW09463;
XX
XX 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding compound peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 1. 14

XX /note= "Preferably linkages are selected from: -

XX CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
XX ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
XX lower alkyl"

XX Modified-site 1

XX /note= "Preferably N-terminus is selected from: -NRR1; -
XX NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
XX benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3
XX substitutions on the phenyl ring selected from lower
XX alkyl, lower alkoxy, chloro, bromo; where R and R1 are
XX independently selected from hydrogen and lower alkyl"

XX Modified-site 14

XX /note= "Preferably C-terminus is -C(O)R2 where R2 is
XX selected from hydroxy, lower alkoxy, and -NR3R4, where R3
XX and R4 are independently selected from hydrogen and lower
XX alkyl, and where the nitrogen atom of the -NR3R4 group
XX can optionally be the amine group of the N-terminus of
XX the peptide forming a cyclic peptide"

XX WO9640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.
XX 07-JUN-1995; 95US-00473604.
XX 07-JUN-1995; 95US-00476168.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00484090.
XX 07-JUN-1995; 95US-00485301.
XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.

XX Claim 18; Page 89; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)
XX receptor (TR). It has a molecular weight of < 8000 Da, and a binding
XX affinity to TR as expressed by an IC50 of no more than about 100 mum. The
XX compound (especially if modified, see features table) can be used for
XX treating patients suffering from haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. The peptide may also be used to maintain the
XX proliferation and growth of TPO-dependent cell lines and for use in
XX biological research, for detecting TPO receptors on living cells

XX Sequence 14 AA;

XX Query Match 100.0%; Score 73; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWLAARA 14

Db 1 IEQPTLRQWLAARA 14

RESULT 2

AAW09468
ID AAW09468 standard; protein; 14 AA.

XX AAW09468;

XX 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding compound peptide (part of a dimer).

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX Key Location/Qualifiers

XX Cross-links 14

XX /note= "Linked to the omega Lys from AAW19534"

XX WO9640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

XX 07-JUN-1995; 95US-00473604.

XX 07-JUN-1995; 95US-00476168.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00484090.

XX 07-JUN-1995; 95US-00485301.

```

XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-051883/05.
XX XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX XX
XX PS Claim 30; Page 91; 106pp; English.
XX CC The present sequence is a compound which binds to thrombopoietin (TPO)
CC receptor (TR). It is part of a dimer linked by the omega amino acid to
CC the omega amino acid in the sequence in AAW19534. The compound can be
CC used for treating patients suffering from haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. The peptide may also be used to maintain the
CC proliferation and growth of TPO-dependent cell lines and for use in
CC biological research, for detecting TPO receptors on living cells
XX SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
DB 1 IEGPTLRQWLAARA 14

RESULT 3
AAW33030
ID AAW33030 standard; peptide; 14 AA.
XX AC AAW33030;
XX XX
XX DT 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX FH Key
XX FT Cross-links 14
XX FT /note= "terminal carboxy group linked to epsilon amino
XX group of Lys16 in AAW33035"
XX PN WO9640750-A1.
XX XX
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX XX
XX PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX XX
XX DR WPI; 1997-052226/05.
XX XX
XX PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Claim 30; Page 91; 106pp; English.
XX CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
XX cell lines
XX SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
DB 1 IEGPTLRQWLAARA 14

RESULT 4
AAW33034
ID AAW33034 standard; peptide; 14 AA.
XX AC AAW33034;
XX XX
XX DT 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX FH Key
XX FT Cross-links 14
XX FT /note= "terminal carboxy group linked to epsilon amino
XX group of Lys16 in AAW33035"
XX PN WO9640750-A1.
XX XX
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX XX
XX PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX XX
XX DR WPI; 1997-052226/05.
XX XX
XX PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Claim 30; Page 91; 106pp; English.
XX CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
XX cell lines
XX SQ Sequence 14 AA;

```

CC mechanism of thrombopoietin signal transduction and receptor activation,
 CC or to maintain the proliferation and growth of thrombopoietin dependent
 CC cell lines
 XX
 SQ Sequence 14 AA;
 Query Match 100.0%; Score 73; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 IEGPTLRQWLAARA 14
 DB 1 IEGPTLRQWLAARA 14
 RESULT 5
 AAW36774
 ID AAW36774 standard; peptide; 14 AA.
 XX
 AC AAW36774;
 XX
 DT 11-MAR-1998 (first entry)
 DE Thrombopoietin receptor binding peptide.
 XX
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 14
 FT /note= "NH2-Ala"
 XX
 PN WO9640750-Al.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US009623.
 XX
 PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX
 DR WPI; 1997-052226/05.
 XX
 PT Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
 XX
 PS Example 9; Page 77; 106pp; English.
 XX
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 SQ Sequence 14 AA;
 Query Match 100.0%; Score 73; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEGPTLRQWLAARA 14
 DB 1 IEGPTLRQWLAARA 14
 RESULT 6
 ADI24843
 ID ADI24843 standard; peptide; 14 AA.
 XX
 AC ADI24843;
 XX
 DT 15-APR-2004 (first entry)
 DE AF 12505 as active moiety for pharmacologically active peptide.
 XX
 KW pharmacologically active peptide conjugate; enzymatic cleavage; pain;
 KW HIV; cancer; diabetes; incontinence; hypertension; amnesia;
 KW Alzheimer's disease; fever; depression; sex hormone regulation;
 KW eating disorder; schizophrenia; osteoporosis; insomnia;
 KW Central nervous system disorder; contraceptive.
 XX
 OS Synthetic.
 XX
 PN WO9946283-Al.
 XX
 PD 16-SEP-1999.
 XX
 PF 09-MAR-1999; 99WO-DK000118.
 XX
 PR 09-MAR-1998; 98DK-00000317.
 XX
 PA (ZEAL-) ZEALAND PHARM AS.
 XX
 PI Larsen BD;
 DR WPI; 1999-561659/47.
 PT New peptide conjugates used for treating, e.g. pain, HIV, depression,
 PT schizophrenia, osteoporosis or insomnia.
 XX
 PS Claim 24; Page 90; 113pp; English.
 XX
 CC The invention relates to a novel pharmacologically active peptide
 CC conjugate having a reduced tendency towards enzymatic cleavage comprises
 CC X and Z, where: (a) X is a pharmacologically active peptide sequence; and
 CC (b) Z is a stabilising peptide sequence of 4-20 amino acid units
 CC covalently bound to X, where each amino acid unit in the stabilizing
 CC peptide sequence, Z being selected from Ala, Leu, Ser, Thr, Tyr, Asn,
 CC Gln, Asp, Glu, Lys, Arg, His, Met, Orn, and amino acid units of formula -
 CC NH-C(R1)(R2)-C(=O)- (I), where: R1 and R2 are H, 1-6C alkyl, phenyl, and
 CC phenyl-methyl, where 1-6C-alkyl is optionally substituted with 1-3
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,
 CC sulfonyl, and carboxy, and phenyl and phenyl-methyl are optionally
 CC substituted with 1-3 substituents selected from 1-6C-alkyl, 2-6C-alkenyl,
 CC halogen, hydroxy, amino, cyano, nitro, sulfonyl, and carboxy, or R1 and R2
 CC 2 together with the C atom to which they are bound form a cyclopentyl,
 CC cyclohexyl or cycloheptyl ring, e.g. 2,4-diaminobutanoic acid and 2,3-
 CC diaminopropanoic acid; the ratio between the half-life of the peptide
 CC conjugate and the half-life of the corresponding active peptide sequence,
 CC X, when treated with carboxypeptidase A or leucine aminopeptidase in
 CC about 50 mM phosphate buffer solution at about pH 7.4 and about 37 deg C
 CC or in serum or plasma is at least about 2 (preferably at least about 10),
 CC or when the pharmacologically active peptide X is not orally absorbed,
 CC the conjugate is adsorbed, or a salt, with the proviso that the
 CC pharmacologically active peptide conjugate is not selected from sequences
 CC (ADI24837)-(ADI24841). The peptide conjugates can be used for treating
 CC e.g. pain, HIV, cancer, diabetes, incontinence, hypertension, amnesia,
 CC Alzheimer's disease, fever, depression, sex hormone regulation, eating
 CC disorders, schizophrenia, osteoporosis or insomnia. They can also be used
 CC for treating e.g. CNS disorders and as contraceptives. The conjugated
 CC peptides are less susceptible to degradation by proteases compared to the
 CC corresponding free pharmacologically active peptides. This sequence
 CC represents a pharmacologically active peptide as the X part of the


```
CC peptide of the invention.
XX
SQ Sequence 14 AA;
  Query Match      100.0%; Score 73; DB 2; Length 14;
  Best Local Similarity 100.0%; Pred. No. 1.6e-05;
  Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IEGPTLRQWLAAARA 14
  |||||
Db 1 IEGPTLRQWLAAARA 14
  |||||

RESULT 7
AAY96515
ID AAY96515 standard; peptide; 14 AA.
AC AAY96515;
XX
DT 04-SEP-2000 (first entry)
DE
XX
XX Thrombopoietin mimetic peptide.
XX
XX Thrombopoietin; mimetic; TPO; TPO; platelet; megakaryocyte; production;
KW anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;
KW immunosuppressive; anti-inflammatory.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 14
FT /note= "subunits in the dimer are covalently bonded at
FT each carboxy terminus through peptide linkage with NH2-
FT (CH2)4-CH(CONH2)-NH-CO-(CH2)2-NH2"
XX
XX WO200024770-A2.
XX
XX 04-MAY-2000.
XX
XX 22-OCT-1999; 99WO-US024834.
XX
XX 23-OCT-1998; 98US-0105348P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Liu C, Feige U, Cheetham J;
XX WPI; 2000-365108/31.
XX
XX Thrombopoietic peptides which activate mpl receptors and increase the
XX production of platelets or platelet precursors, useful for treatment of
XX diseases which involve thrombocytopenia.
XX
XX Claim 7; Page 60; 91pp; English.
XX
XX A compound which binds to an mpl receptor comprising a thrombopoietin
XX mimetic peptide (TMP) dimer joined by a linker [TMP 1-(L1)-TMP 2], is
XX new. TMP 1 and TMP 2 are amino acid sequences varying from at least 10 to
XX 14 residues in length comprising X 2-X 1 0, X 2-X 1 1, X 2-X 1 2, X 2-
XX X 1 3, X 2-X 1 4, X 1-X 1 0, X 1-X 1 1, X 1-X 1 2, X 1-X 1 3, and X 1-
XX X 1 4. X 1 = I, A, V, L, S or R; X 2 = E, D, K or V; X 3 = G or A; X 4 =
XX P; X 5 = T or S; X 6 = L, I, V, A or F; X 7 = R or K; X 8 = Q, N, or E;
XX X 9 = W, Y or F; X 1 0 = L, I, V, A, F, M, or K; X 1 1 = A, I, V, L, F,
XX S, T, K, H, or E; X 1 2 = A, I, V, L, F, G, S, or Q; X 1 3 = R, K, T, V,
XX N, or G; X 1 4 = A, I, V, L, F, T, R, E, or G; L 1 = linker comprising
XX 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate
XX the c-Mpl receptor which mediates the activity of endogenous
XX thrombopoietin. The TMPs are useful for increasing the production of
XX platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which
XX is useful for treatment of diseases which involve thrombocytopenia, e.g.
XX aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency
XX virus associated ITP, and systemic lupus erythematosus
XX
```

```
SQ Sequence 14 AA;
  Query Match      100.0%; Score 73; DB 3; Length 14;
  Best Local Similarity 100.0%; Pred. No. 1.6e-05;
  Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IEGPTLRQWLAAARA 14
  |||||
Db 1 IEGPTLRQWLAAARA 14
  |||||

RESULT 8
AAB16962
ID AAB16962 standard; peptide; 14 AA.
XX
AC AAB16962;
XX
DT 31-OCT-2000 (first entry)
DE
XX
XX TPO-mimetic peptide TMP SEQ ID NO:13.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
XX Synthetic.
XX
XX WO200024782-A2.
XX
XX 04-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US025044.
XX
XX 23-OCT-1998; 98US-0105371P.
XX
XX 22-OCT-1999; 99US-00428082.
XX
XX (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 19; Page 189; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX (L2)d-P2-(L3)e-P'3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX P3, and P4 = are each independently sequences of pharmacologically active
XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX cells from the present invention can be used for producing pharmaceutical
XX compositions. The compositions are useful for treating cancer, asthma,
XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX such as Fc receptor binding, protein A binding, complement fixation, and
XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX AAB18003 represent nucleotide and amino acid sequences used in the
XX exemplification of the present invention
XX
XX Sequence 14 AA;
  Query Match      100.0%; Score 73; DB 3; Length 14;
```

Best Local Similarity 100.0%; Pred. No. 1.6e-05; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14

RESULT 9
 AAU25827
 ID AAU25827 standard; peptide; 14 AA.
 AC AAU25827;
 DT 17-DEC-2001 (first entry)
 DE Human thrombopoietin receptor (TPO-R) activator peptide #13.
 DX
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 OS Homo sapiens.
 XX
 PN US6251864-B1.
 XX
 PD 26-JUN-2001.
 XX
 PF 01-MAR-2000; 2000US-00516704.
 XX
 PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US0009623.
 PR 15-AUG-1996; 96US-00699027.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 DR WPI; 2001-564142/63.
 XX
 PT Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 PS Disclosure; Col 69-70; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

Sequence 14 AA;

Query Match 100.0%; Score 73; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14

RESULT 10
 AAU26004
 ID AAU26004 standard; peptide; 14 AA.
 XX
 AC AAU26004;
 XX
 DT 17-DEC-2001 (first entry)
 DE Human thrombopoietin receptor (TPO-R) activator peptide #190.
 DX
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 OS Homo sapiens.
 XX
 PN US6251864-B1.
 XX
 PD 26-JUN-2001.
 XX
 PF 01-MAR-2000; 2000US-00516704.
 XX
 PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US0009623.
 PR 15-AUG-1996; 96US-00699027.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 DR WPI; 2001-564142/63.
 XX
 PT Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 PS Disclosure; Col 147; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

Sequence 14 AA;

Query Match 100.0%; Score 73; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
| | | | | | | | | | | | | |
DB 1 IEGPTLRQWLAAARA 14

RESULT 11

ID ABB72853 standard; peptide; 14 AA.

XX

AC ABB72853;

XX

DT 05-APR-2002 (first entry)

XX

DE TPO mimetic peptide SEQ ID NO:13.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
KW TNP mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytosolic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

XX

PD 08-NOV-2001.

XX

PF 02-MAY-2001; 2001WO-US014310.

XX

PR 03-MAY-2000; 2000US-00563286.

XX

PA (AMGE-) AMGEN INC.

XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX

DR WPI; 2002-130313/17.

XX

XX Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX

PS Claim 39; Page 43; 176pp; English.

XX

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 5; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.6e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14

| | | | | | | | | | | | | |
DB 1 IEGPTLRQWLAAARA 14

RESULT 12

ABP51669

ID ABP51669 standard; peptide; 14 AA.

XX

AC ABP51669;

XX

DT 01-OCT-2002 (first entry)

XX

DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:1.

XX

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;

KW complementarity determining region; immunoglobulin; antianaemic;

KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX Homo sapiens.

OS Synthetic.

XX WO200246238-A2.

XX

PN 13-JUN-2002.

XX

PD 05-DEC-2001; 2001WO-US047656.

XX

PF 05-DEC-2000; 2000US-0251448P.

XX

PR 04-MAY-2001; 2001US-0288889P.

XX

PR 29-MAY-2001; 2001US-0294068P.

XX

PA (ALEX-) ALEXION PHARM INC.

XX

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

PI WPI; 2002-566610/60.

XX

XX A novel immunogen molecule comprising a region in which amino acid

PT residues corresponding to at least a portion of the complementary

PT determining region are replaced or fused with an erythropoietin or

PT thrombopoietin mimetic.

XX

PS Claim 18; Page 6; 113pp; English.

XX

XX The present invention describes an immunoglobulin molecule or its fragment

CC (I) comprising a region where amino acid residues corresponding to at

CC least a portion of the complementary determining region (CDR) are

CC replaced or fused with biologically active peptides e.g. a peptide

CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,

CC that is flanked with proline at its carboxy terminus. (I) has

CC antianaemic, haemostatic and nephrotropic activities, and can be used as

CC a stimulator of proliferation, differentiation and maturation of

CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful

CC for stimulating proliferation, differentiation or growth of

CC promegakaryocytes or megakaryocytes, where (I) is contacted with

CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or

CC more of its CDRs fused to an EPO mimetic, is useful for increasing the

CC production of red blood cells, where (I) is contacted with haematopoietic

CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiencies in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 14 AA;
 SQ Query Match 100.0%; Score 73; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 Db |||||K|||||
 1 IEPTLRQWLAARA 14

RESULT 13
 ID AAE18011 standard; peptide; 14 AA.

XX AAE18011;
 AC 07-MAY-2002 (first entry)

XX Human ligand #31 attached to an adenoviral vector.

XX Human; adenoviral coat protein; non-native ligand; cell-surface receptor;
 KW therapy; anti-tumour agent; tumour necrosis factor; cancer; brain; lung;
 KW ovary; breast; prostate.

XX Homo sapiens.

XX WO200192549-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017391.

XX 31-MAY-2000; 2000US-0208451P.

XX 02-AUG-2000; 2000US-00631191.

XX (GENV-) GENVEC INC.

XX Wickham TJ, Kovessdi I, Roelvink PW, Einfeld D, Brough DE;
 PI Lizonova A;

XX WPI; 2002-147620/19.

XX Adenoviral coat protein which permits production of adenoviral vectors
 PT that bind and infect host cells not naturally infected by adenovirus,
 PT comprises various non-native ligands.

XX Claim 4; Page 45; 45pp; English.

XX The invention relates to adenoviral coat proteins comprising various non-
 CC native ligands. The invention provides a method of controlled gene
 CC expression utilising selectively replication competence and also a method
 CC and a composition for targeting an adenoviral vector. A system
 CC comprising a cell having a non-native cell-surface receptor, and a virus
 CC having a non-native ligand which binds the non-native cell-surface
 CC receptor of the cell is useful for propagating a virus and also for
 CC assaying gene function. The system is also useful for isolating a nucleic
 CC acid encoding a product comprising a desired property. Further the system
 CC is useful for identifying functionally related coding sequences.

XX Adenoviral vector comprising a non-native nucleic acid encoding a
 CC therapeutic agent such as anti-tumour agent, preferably tumour necrosis
 CC factor and a second non-native nucleic acid encoding an agent that
 CC facilitates imaging and a targeting agent is useful for treating an
 CC animal. The therapeutic agent can be used to treat cancer of the brain,
 CC lung, ovary, breast and prostate. The present sequence is human non-
 CC native ligand attached to an adenoviral vector

XX Sequence 14 AA;
 SQ Query Match 100.0%; Score 73; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 Db |||||K|||||
 1 IEPTLRQWLAARA 14

RESULT 14

ID ABG71747 standard; peptide; 14 AA.

XX ABG71747;

XX 20-JAN-2003 (first entry)

XX TPO receptor, MPL, agonist peptide, AFI2505.

XX Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
 KW complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
 KW thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
 KW T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
 KW proliferation; growth; differentiation; haematopoietic cell;
 KW platelet progenitor cell; immune disorder; thrombocytopenia;
 KW disseminated intravascular coagulation; stem cell; transplantation;
 KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.

XX Synthetic.

XX WO200278612-A2.

XX 10-OCT-2002.

XX 02-APR-2002; 2002WO-US010301.

XX 02-APR-2001; 2001US-0281183P.

XX (PURD) PURDUE PHARMA LP.

XX Soltis DA, Burch RM, Ogert RA;

XX WPI; 2003-040615/03.

XX New thrombopoietin synthebodies, useful for stimulating proliferation,
 PT growth, or differentiation of hematopoietic cells, for treating or
 PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.

XX Claim 62; Page 71; 97pp; English.

XX The invention discloses a variant of an immunoglobulin (Ig) variable
 CC heavy or light chain domain that comprises at least one complementarity
 CC determining region (CDR) and framework regions flanking the CDR. The CDR
 CC also has added or substituted to it, at least one binding sequence which
 CC is heterologous to the CDR and is an antigenic, agonistic sequence from a
 CC thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
 CC can be a binding sequence heterologous to the CDR, a cytotoxic T-
 CC lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
 CC cell sequence or a combination of each. The variant or thrombopoietin
 CC synthebody, pharmaceutical and vaccine compositions are useful for
 CC stimulating proliferation, growth or differentiation of haematopoietic
 CC cells, particularly platelet progenitor cells. The variants are also
 CC useful for treating or preventing haematopoietic or immune disorders
 CC resulting from chemotherapy, radiation therapy, or bone marrow
 CC transfusions (e.g. thrombocytopenia or disseminated intravascular
 CC coagulation). Compositions comprising the synthebodies can be used for
 CC the mobilisation, amplification and ex vivo expansion of stem cells and
 CC committed precursor cells for autologous and allogeneic transplantation
 CC as well as for the expansion of stem cells for gene therapy. They are
 CC also useful as diagnostic or analytical reagents for studying the

CC function of thrombopoietin and its receptor in vivo or in vitro. The
CC sequence presented is the TPO receptor (MPL) agonist peptide, AF12505
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
|||||
Db 1 IEGPTLRQWLAAARA 14

RESULT 15

ABR62907
ID ABR62907 standard; peptide; 14 AA.
XX
AC ABR62907;
XX
DT 04-DEC-2003 (first entry)
XX
DE Thrombopoietin mimetic peptide AF12505.
XX
KW Thrombopoietin; mimetic; thrombocytopaenia; antibody targeting.
XX
OS Synthetic.
XX
PN WO2003059251-A2.
XX
PD 24-JUL-2003.
XX
PF 22-OCT-2002; 2002WO-US033991.
XX
PR 22-OCT-2001; 2001US-0344614P.
XX
PA 19-SEP-2002; 2002US-0412455P.
XX
PI (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Rader C, Sinha SC, Lerner R;
XX
WPI; 2003-636673/60.
XX
XX
XX Antibody targeting compound useful e.g. for diagnostic immunoassays and
XX treating microbial diseases comprises targeting or biological agent
XX covalently linked to combining site of the antibody.
XX
XX Example 7; Page 62; 56pp; English.

The present sequence is that of thrombopoietin (TPO) mimetic peptide
AF12505, which mimics the activity of recombinant TPO. The invention
provides antibody targeting compounds that are used to reprogram the
specificity of an antibody. The antibody targeting compound is linked to
the combining site of the antibody, such that the modified antibody takes
on the binding specificity of the targeting agent. In an example from the
invention, a TPO receptor targeting antibody compound was prepared by
covalently linking peptide ABL12505 to aldolase monoclonal antibody 38C2.
The TPO receptor targeting antibody compound can be used to treat
CC thrombocytopaenia resulting from chemotherapy and bone marrow
CC transplantation
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
|||||
Db 1 IEGPTLRQWLAAARA 14

RESULT 16

ADC33697
ID ADC33697 standard; peptide; 14 AA.
XX
AC ADC33697;
XX
DT 18-DEC-2003 (first entry)
XX
DE Erythropoietin receptor/erythropoietin consensus peptide SEQ ID NO.1.
XX
KW chimeric retrovirus envelope protein; ecotropic envelope protein;
KW cytotstatic; gene therapy; cancer.
XX
OS Synthetic.
XX
PN WO2003076596-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US007323.
XX
PR 08-MAR-2002; 2002US-0362655P.
XX
PA (UTMA-) UNIV MASSACHUSETTS.
XX
PI Green MR, Gollan TJ;
XX
WPI; 2003-722332/68.
XX
XX New chimeric retrovirus envelope protein comprising an ecotropic envelope
XX protein and a heterologous short peptide ligand inserted within the
XX ecotropic envelope protein useful for treating cancer.
XX
XX Disclosure; SEQ ID NO 1; 42pp; English.

The present invention describes a chimeric retrovirus envelope protein
(I) comprising an ecotropic envelope protein and a heterologous short
peptide ligand inserted within the ecotropic envelope protein. Also
described: (1) a nucleic acid molecule comprising a sequence encoding the
recombinant chimeric envelope protein; (2) a vector comprising a nucleic
acid sequence encoding the chimeric envelope protein; (3) a recombinant
retroviral particle comprising a chimeric envelope protein comprising a
heterologous short peptide ligand; (3) altering retroviral tropism; (4)
identifying a nucleic acid sequence encoding the chimeric envelope
protein that alters viral tropism; (5) delivering a nucleic acid sequence
to a cell; and (6) treating cancer. (1) has cytostatic activity and can
be used in gene therapy. The chimeric retrovirus envelope protein is
useful for treating cancer, which comprises providing a cancer cell, e.g.
human cancer cell and infecting the cancer cell with a virus, e.g.
retrovirus comprising the chimeric envelope protein comprising a
heterologous short peptide ligand and a therapeutically useful gene, e.g.
encoding thymidine kinase. The present sequence represents an
erythropoietin receptor/erythropoietin consensus peptide, which is given
in the exemplification of the present invention.

Sequence 14 AA;

Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
|||||
Db 1 IEGPTLRQWLAAARA 14

RESULT 17

ADN59652
ID ADN59652 standard; peptide; 14 AA.
XX
AC ADN59652;
XX
DT 01-JUL-2004 (first entry)
XX

DS Thrombopoietin mimetic peptide (TMP), seq id 1.
XX Haemostatic; antianemic; immunosuppressive; platelet;
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;
KW thrombocytopenia; aplastic anaemia; autoimmune thrombocytopenia;
KW autoimmune haemolytic anaemia; Hughes's syndrome;
KW lupoid thrombocytopenia.
XX
OS Homo sapiens.
XX
XX WO2003031589-A2.
XX
XX 17-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032552.
XX
XX 11-OCT-2001; 2001US-0328666P.
XX
XX 10-OCT-2002; 2002US-00269806.
XX
XX (AMGE-) AMGEN INC.
XX
XX Min H, Sitney KC, Hartley C;
XX WPI; 2003-403101/38.
XX
XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
PT which stimulate the production of platelets and/or the production of
PT platelet precursors, useful for treating thrombocytopenia.
XX
XX Disclosure; SEQ ID NO 1; 126pp; English.
XX
XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
XX binds to the c-mpl (mpl) receptor, and which stimulates the production of
XX platelets and/or the production of platelet precursors, is new. Further
XX disclosed is a composition of matter (II) that binds to an mpl receptor,
XX and a pharmaceutical composition comprising (II) and a carrier. The
XX pharmaceutical composition of the invention is useful for treating
XX thrombocytopenia in an animal, and for increasing megakaryocytes or
XX platelets in a patient. The TMP of the invention is useful for treating
XX conditions involving a megakaryocyte and/or platelet deficiency, e.g.
XX disease conditions involving thrombocytopenia such as aplastic anaemia,
XX autoimmune thrombocytopenia, drug induced immune thrombocytopenia,
XX autoimmune haemolytic anaemia, Hughes's syndrome and lupoid
XX thrombocytopenia. The TMP of the invention is also useful for
XX maintaining the viability or storage life of platelets and/or
XX megakaryocytes and its derived cells. The compounds demonstrate an
XX improved ability to bind to and/or trigger transmembrane signal through,
XX i.e. activating, the mpl receptor the compounds have superior
XX thrombopoietic activity, i.e. the ability to stimulate, in vivo and in
XX vitro, the production of platelets and/or megakaryocytic activity,
XX i.e. the ability to stimulate, in vivo and in vitro, the production of
XX platelet precursors. Further, certain of the compounds also exhibit
XX superior therapeutic properties, such as improved plasma half-life,
XX biological activity and in vivo circulation time. The current sequence
XX represents a TMP of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
RESULT 18
ADL27293
ID ADL27293 standard; protein; 14 AA.
XX
XX ADL27293;
AC

XX 03-JUN-2004 (first entry)
XX
XX Amino acid sequence of a thrombopoietin agonist peptide.
XX fusion protein; C4bp; alpha chain; systemic lupus erythematosus.
XX
XX Homo sapiens.
XX Synthetic.
XX WO2004020639-A2.
XX
XX 11-MAR-2004.
XX
XX 12-AUG-2003; 2003WO-EP008928.
XX
XX 14-AUG-2002; 2002EP-00292043.
XX
XX (AVID-) AVIDIS SA.
XX
XX Garnier L, Hill F, Julien M;
XX WPI; 2004-239202/22.
XX
XX Obtaining a recombinant fusion protein, useful for treating lupus,
XX comprises providing a prokaryotic host cell carrying a nucleic acid
XX encoding the recombinant protein operably linked to a promoter functional
XX in the prokaryotic cell.
XX
XX Claim 8; Page 48; 69pp; English.
XX
XX The specification describes a method for obtaining a recombinant fusion
XX protein comprising a scaffold of a C-terminal core protein of C4bp alpha
XX chain, where the recombinant fusion protein is capable of forming
XX multimers in soluble form in a prokaryotic host cell. The method
XX comprises providing a prokaryotic host cell carrying a nucleic acid
XX encoding the recombinant protein operably linked to a promoter functional
XX in the prokaryotic cell, culturing the host cell under conditions where
XX the recombinant protein is expressed, and recovering the recombinant
XX protein where the protein is recovered in multimeric form without
XX performing a scaffold refolding step. The protein is useful for treating
XX systemic lupus erythematosus. The present sequence represents a
XX thrombopoietin agonist peptide, which is used to produce fusion proteins
XX of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
RESULT 19
ADM72483
ID ADM72483 standard; peptide; 14 AA.
XX
XX ADM72483;
XX
XX 17-JUN-2004 (first entry)
XX
XX TPO mimetic peptide fragment.
XX
XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
XX stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
XX Synthetic.
XX
XX WO2004026332-A1.
XX

PD 01-APR-2004.
 XX 18-SEP-2003; 2003WO-US029701.
 XX 18-SEP-2002; 2002US-0411700P.
 PR 18-SEP-2002; 2002US-0411779P.
 XX (THRE-) 3-DIMENSIONAL PHARM INC.
 PA Kaushansky K, Macdonald BR;
 XX WPI; 2004-283153/26.
 XX Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX Disclosure; Fig 2; 32pp; English.
 PS The invention relates to a method (M1) for increasing haematopoietic stem
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing haematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 100.0%; Score 73; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14
 |||||
 RESULT 20
 ADQ16584
 ID ADQ16584 standard; peptide; 14 AA.
 XX
 AC ADQ16584;
 XX
 XX 09-SEP-2004 (first entry)
 DT
 Query Match 100.0%; Score 73; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14
 |||||
 RESULT 21
 ADT92482
 ID ADT92482 standard; peptide; 14 AA.
 XX
 AC ADT92482;
 XX
 XX 13-JAN-2005 (first entry)
 DT
 DE Thrombopoietin mimetic peptide AFI2505 SeqID 5.
 XX
 KW integrin targeting compound; RGD peptidomimetic; integrin-alpha;
 KW angiogenesis; bone metabolism; inflammation; cancer; cell growth;
 KW antiangiogenic; osteopathic; antiinflammatory; cytostatic; TPO;
 KW thrombopoietin.
 XX
 OS Unidentified.
 OS
 XX WO2004091542-A2.
 PN
 XX 28-OCT-2004.
 PD
 XX 15-APR-2004; 2004WO-US012034.
 PF

XX Agonist TPO mimetic peptide SEQ ID NO:1.
 DE immunoglobulin; complementarity determining region; CDR; peptide mimetic;
 KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
 KW immunotherapy; thrombocytopenia.
 XX
 OS Unidentified.
 OS
 PN WO2004050017-A2.
 XX
 XX 17-JUN-2004.
 PD
 XX 17-NOV-2003; 2003WO-US036894.
 PF
 XX 02-DEC-2002; 2002US-00307724.
 PR
 XX (ALEX-) ALEXION PHARM INC.
 XX
 PA Bowdish KS, Frederickson S, Renshaw M;
 PI WPI; 2004-460973/43.
 XX
 DR New immunoglobulin molecule comprising a region, where two
 XX complementarity determining regions (CDRs) are replaced with EPO mimetic
 PT or a TPO mimetic, useful for treating thrombocytopenia.
 PT
 XX Claim 8; SEQ ID NO 1; 107pp; English.
 PS The invention relates to a novel immunoglobulin molecule or its fragment
 CC comprising a region where amino acid residues corresponding to at least a
 CC portion of a two complementarity determining regions (CDRs) are replaced
 CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
 CC invention has immunosuppressive activity, and may have a use in
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
 CC The present sequence represents a TPO mimetic peptide.
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 100.0%; Score 73; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14
 |||||
 RESULT 21
 ADT92482
 ID ADT92482 standard; peptide; 14 AA.
 XX
 AC ADT92482;
 XX
 XX 13-JAN-2005 (first entry)
 DT
 DE Thrombopoietin mimetic peptide AFI2505 SeqID 5.
 XX
 KW integrin targeting compound; RGD peptidomimetic; integrin-alpha;
 KW angiogenesis; bone metabolism; inflammation; cancer; cell growth;
 KW antiangiogenic; osteopathic; antiinflammatory; cytostatic; TPO;
 KW thrombopoietin.
 XX
 OS Unidentified.
 OS
 XX WO2004091542-A2.
 PN
 XX 28-OCT-2004.
 PD
 XX 15-APR-2004; 2004WO-US012034.
 PF

XX 15-APR-2003; 2003US-0463456P.
PR 30-SEP-2003; 2003US-0507887P.
XX (COVX-) COVX PHARM INC.
XX Tamiz A, Bradshaw CW;
PI WPI; 2004-784488/77.
XX
DR
XX
XX New nitrogen containing integrin targeting compounds are integrin alpha-v
PT beta-3 antagonists useful for the treatment or prevention of e.g. bone
PT metabolism, inflammation, cancer and cell growth.
XX
XX Example 7; SEQ ID NO 5; 182pp; English.
XX
XX This invention relates to novel nitrogen containing integrin targeting
CC compounds, their stereoisomers, tautomers or salts thereof. Specifically,
CC it refers to integrin targeting compounds that comprise small MW, RGD
CC peptidomimetic integrin targeting agent-linker conjugates that are linked
CC to polymers such as a protein. The present invention describes compounds
CC that can be useful for the treatment or prevention of diseases that
CC involve integrin-alpha, and can be used to treat a defect in
CC angiogenesis, bone metabolism, inflammation, cancer or cell growth.
CC Accordingly, they exhibit antiangiogenic, osteopathic, antiinflammatory
CC and cytostatic activities. This peptide sequence is a TPO mimetic peptide
CC given in an exemplification of the invention.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IEPTLRQWLAARA 14
Db 1 IEPTLRQWLAARA 14
RESULT 22
ADU70210
ID ADU70210 standard; peptide; 14 AA.
XX
AC ADU70210;
XX
XX 10-FEB-2005 (first entry)
XX
XX Thrombopoietin receptor modulator peptide #5.
KW hematological disease; drug delivery; thrombopoietin receptor modulator;
KW diagnosis; pharmaceutical.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /label= OTHER
FT /note= "Covalently linked to a spacer/PEG moiety which is
FT in turn lionked to another copy of this sequence forming
FT a dimer"
FT Modified-site 14 /note= "Ala is amidated"
FT
XX WO2004100997-A2.
XX
XX 25-NOV-2004.
XX
XX 12-MAY-2004; 2004WO-US014887.
XX
XX 12-MAY-2003; 2003US-0469996P.
XX (AFFY-) AFFYMAX INC.
XX

PI Holmes C, Yin K, Tumelty D;
XX WPI; 2005-039305/04.
XX
XX New peptide-based compound comprising a peptide moiety, a spacer moiety
PT and a water-soluble polymer moiety useful as pharmaceuticals.
XX
XX Example 15; Page 46; 57pp; English.
XX
XX The invention relates to a peptide-based compound (I) comprising a
CC peptide moiety, a spacer moiety and a water-soluble polymer moiety (e.g.
CC polyethylene glycol, PEG). Also included is a pharmaceutical composition
CC comprising (I) and one or more pharmaceutically acceptable diluents,
CC preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. The
CC peptide moiety comprises one or more peptides that bind to erythropoietin
CC -receptors or thrombopoietin-receptors. The peptide based compound is
CC useful as a pharmaceutical. The peptide moiety particularly
CC erythropoietin receptors are essential in the process of red blood cell
CC formation and has useful applications in both diagnosis and the treatment
CC of blood disorders characterized by low or defective red blood cell
CC production. The peptide based compound has improved antigenicity and
CC durability. Polyethylene glycolylation leads to improved drug efficacy and
CC safety, increases bioavailability and decreases immunogenicity and dosing
CC frequency. The present sequence is a thrombopoietin-receptor modulating
CC peptide.
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IEPTLRQWLAARA 14
Db 1 IEPTLRQWLAARA 14
RESULT 23
ADU75982
ID ADU75982 standard; peptide; 14 AA.
XX
AC ADU75982;
XX
XX 10-FEB-2005 (first entry)
XX
XX Peptide-based compound associated peptide #3.
XX
XX antianemic; pharmaceutical; blood; blood disorder.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /label= OTHER
FT /note= "Two of these molecules are optionally
FT attached by the N-terminal by a spacer molecule that can
FT undergo PGylation"
XX WO2004101600-A2.
XX
XX 25-NOV-2004.
XX
XX 12-MAY-2004; 2004WO-US014888.
XX
XX 12-MAY-2003; 2003US-0470246P.
XX (AFFY-) AFFYMAX INC.
XX
XX Yin K, Holmes C, Lalonde G, Balu P, Tumelty D, Schatz P;
XX WPI; 2005-039327/04.
XX
XX Peptide-based compound useful as e.g. pharmaceutical for the treatment of
PT

PT blood disorders comprises a peptide moiety and a linear
PT poly(ethyleneglycol) moiety.
XX
XX Disclosure; Page 47; 56pp; English.
XX
XX The invention describes a peptide-based compound (A) comprising a peptide
XX moiety (B) and a linear poly(ethyleneglycol) moiety (C) having molecular
XX weight of more than 20 kDaltons. The compounds are useful as
XX pharmaceuticals for the treatment of blood disorders characterized by low
XX as peptide modifying reagent. The compounds exhibit improved antigenicity
XX and durability. This is the amino acid sequence of peptide that may be
XX present in the peptide moieties of the invention.
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 73; DB 9; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 IEGPTLRQWLAARA 14
XX |||||||
XX Db 1 IEGPTLRQWLAARA 14
XX
XX RESULT 24
XX ADV44319
XX ID ADV44319 standard; peptide; 14 AA.
XX AC
XX AC ADV44319;
XX
XX DT 10-MAR-2005 (first entry)
XX
XX DE Agonist TPO mimetic peptide seqid 1.
XX
XX KW anti-HIV; cytostatic; gene therapy; antibody engineering; diabetes;
XX HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
XX transplant rejection; prophylaxis; myeloproliferative disorder;
XX hematological disease; TPO mimetic peptide.
XX
XX OS Unidentified.
XX
XX PN WO2004108078-A2.
XX
XX PD 16-DEC-2004.
XX
XX PF 26-MAY-2004; 2004WO-US016574.
XX
XX PR 02-JUN-2003; 2003US-00452590.
XX
XX PA (ALEX-) ALEXION PHARM INC.
XX
XX PI Bowdish KS, Frederickson S, Renshaw M, Orenicia C;
XX WPI; 2005-031598/03.
XX
XX New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
XX Claim 8; SEQ ID NO 1; 139pp; English.
XX
XX The invention describes an immunoglobulin molecule or its fragment
XX comprising a region where amino acid residues corresponding to at least
XX a portion of two CDRs are replaced with a peptide mimetic selected from
XX an EPO mimetic or a TPO mimetic, or a region where amino acid residues
XX corresponding to at least a portion of a CDR is replaced by a peptide
XX mimetic including SEQ ID NO. 126 (not defined in the specification),
XX where X at each occurrence represents any amino acid. Also described are:
XX a nucleic acid encoding an immunoglobulin molecule or its fragment; an
XX expression vector comprising the nucleic acid of (1); a host cell
XX transformed with the expression vector of (2); producing an

CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 73; DB 9; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 IEGPTLRQWLAARA 14
XX |||||||
XX Db 1 IEGPTLRQWLAARA 14
XX
XX RESULT 25
XX AEB12792
XX ID AEB12792 standard; peptide; 14 AA.
XX AC AEB12792;
XX
XX DT 08-SEP-2005 (first entry)
XX
XX DE TPO mimetic peptide, SEQ ID 1.
XX
XX KW TPO; Thrombopoietin; protein therapy; antibody engineering;
XX hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
XX cardiac failure; diabetes; obesity.
XX
XX OS Synthetic.
XX
XX PN WO2005060642-A2.
XX
XX PD 07-JUL-2005.
XX
XX PF 15-DEC-2004; 2004WO-US041946.
XX
XX PR 15-DEC-2003; 2003US-00737290.
XX
XX PA (ALEX-) ALEXION PHARM INC.
XX
XX PI Bowdish KS, Frederickson S, Renshaw M, Orenicia C;
XX WPI; 2005-479402/48.
XX
XX New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
XX Example 1; SEQ ID NO 1; 152pp; English.
XX
XX The invention relates to an immunoglobulin (Ig) molecule or its fragment
XX comprising a region where amino acid residues corresponding to at least a
XX portion of a complementarity determining regions (CDR) is replaced with a
XX peptide selected from human brain natriuretic protein (hBNP), hBNP
XX mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
XX mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
XX PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
XX molecule, an expression vector comprising the nucleic acid, a host cell
XX transformed with the expression vector, producing an immunoglobulin
XX molecule (or its fragment, comprising the host cell under
XX conditions suitable for expression of the immunoglobulin or its
XX fragment), a composition comprising an immunoglobulin (or its fragment)
XX and a pharmaceutically acceptable carrier, treating congestive heart

CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBMP and hBMP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
CC glucagons, glucagon mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1),
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to
CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a thrombopoietin (TPO) mimetic, erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 73; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.6e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

Qy 1 IEGPTLRQWLAAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAAARA 14

RESULT 26
AAW35416
ID AAW35416 standard; peptide; 15 AA.

XX AC AAW35416;

XX DT 11-MAR-1998 (first entry)

XX DE Thrombopoietin receptor binding peptide.

XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
XX KW haematological disorder; thrombocytopaenia; chemotherapy;
XX KW radiation therapy; bone marrow transfusion; diagnosis;
XX KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical
FT peptide"

FT Modified-site 15 /note= "NH2-Ala"

XX FN WO9640750-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WFI; 1997-052226/05.

XX PT Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.

XX PS Example 9; Page 73; 106pp; English.

XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX CC used to treat disorders which are susceptible to treatment with a
XX CC thrombopoietin agonist, preferably haematological disorders and
XX CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
XX CC marrow transfusions. It can also be used diagnostically, e.g. to
XX CC investigate the mechanism of thrombopoietin signal transduction and
XX CC receptor activation, or to maintain the proliferation and growth of
XX CC thrombopoietin dependent cell lines

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
| | | | | | | | | |
Db 2 IEGPTLRQWLAAARA 15

RESULT 27
AAW36776

ID AAW36776 standard; peptide; 15 AA.

XX AC AAW36776;

XX DT 11-MAR-1998 (first entry)

XX DE Thrombopoietin receptor binding peptide.

XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
XX KW haematological disorder; thrombocytopaenia; chemotherapy;
XX KW radiation therapy; bone marrow transfusion; diagnosis;
XX KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Cross-links 1 /note= "linked via disulfide bond to Cys1 of identical
FT peptide"

FT Modified-site 15 /note= "NH2-Ala"

XX FN WO9640750-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-052226/05.
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX
 XX Example 9; Page 77; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 73; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARA 14
 DB 2 IEGPTLRQWLAAARA 15
 RESULT 28
 AA66712
 ID AA66712 standard; peptide; 15 AA.
 XX
 AC AA66712;
 XX
 DT 01-DEC-1998 (first entry)
 XX
 DE Peptide chain of compound which binds to the thrombopoietin receptor.
 XX
 KW thrombopoietin receptor; haematological disorder; screening; agonist;
 KW assay; megakaryocyte; blood disorder; thrombocytopenia; TPO.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..14
 FT Modified-site 15 /note= "thrombopoietin receptor agonist"
 FT /note= "Epsilon amino group of Lys, in its amide form, is
 FT attached to another peptide chain identical to the region
 FT (residues 1 to 14) of this peptide"
 XX
 PN W09825965-A2.
 XX
 PD 18-JUN-1998.
 XX
 PF 09-DEC-1997; 97WO-BF006850.
 XX
 PR 11-DEC-1996; 96US-00764640.
 XX
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
 PI Yin Q;
 XX
 XX WPI; 1998-377261/32.
 XX
 XX New peptide compound(s) which can bind and activate thrombopoietin
 PT receptor - may be used in treating haematological disorders and in
 PT methods for screening for new thrombopoietin receptor agonists.
 PT
 XX

PS Claim 2; Page 60; 78pp; English.
 XX
 CC The invention relates to peptide compounds composed of two peptide chains
 CC attached to each of the amino groups of a single lys in the amide form.
 CC The compounds are of formula (Pep1)(Pep2)K(NH2), where Pep1 is of
 CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9-X10; and pep2 is of
 CC formula: X1-I-E-X2-P-T-L-X3-X4-X5-L-X6-X7-X8-X9'-X10'. X1 = H or acyl; X2
 CC = Gly or Sar (sarcosine); X3 = Arg, Ala, Nle (norleucine) or N-Phe; X6 =
 CC acetyllysine; X4 = Gln or Glu; X5 = Trp, L-1-naphthylalanine or Phe; X6 =
 CC Ala, 5-aminopentanoic acid or 2-aminobutyric acid; X7 = Ala,
 CC diphenylalanine, or is absent; X8 = Arg, p- amino-phenylalanine, N-
 CC acetyl-lysine, or is absent; X9, X9' = Ala, beta Ala, N-methyl-alanine,
 CC Sar, or is absent; X10, X10' = beta Ala or is absent. The new peptides
 CC are capable of binding to, and activating, the thrombopoietin (TPO)
 CC receptor. They may be used in vitro as tools for understanding the
 CC biological role of TPO. They may be used as competitive binders in assays
 CC to screen for new TPO receptor agonists. They may be used as reagents for
 CC detecting TPO receptors in living cells, biological fluids, etc. They may
 CC be used to maintain growth and proliferation of TPO-dependent cells and
 CC for in vitro expansion of megakaryocytes. They may be used to activate
 CC TPO receptors in vivo, e.g., to treat blood disorders or
 CC thrombocytopenia associated with bone marrow transfusions, radiotherapy
 CC or chemotherapy. The present sequence represents a specific example of
 CC (Pep1)K(NH2)
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 73; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IEGPTLRQWLAAARA 14
 DB 1 IEGPTLRQWLAAARA 14
 RESULT 29
 AA620684
 ID AA620684 standard; peptide; 15 AA.
 XX
 AC AA620684;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE Thrombocyte generation promoting peptide.
 XX
 KW Thrombocyte; promotion; generation.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 15
 FT /note= "optionally amidated; optionally attached to the C
 FT -terminal cysteine of a similar peptide"
 XX
 PN CN1254718-A.
 XX
 PD 31-MAY-2000.
 XX
 PF 20-NOV-1998; 98CN-00125011.
 XX
 PR 20-NOV-1998; 98CN-00125011.
 XX
 PA (BIOL-) INST BIOLOGICAL ENG CHINESE ACAD MILITAR.
 XX
 PI Cheng D, Li C, Huang P;
 XX
 DR WPI; 2000-533568/49.
 XX
 PT Active peptide.
 XX
 XX Claim 1; Page 1; 5pp; Chinese.
 XX

CC The present invention discloses an active peptide which promotes
 CC thrombocyte generation. The active peptide can be synthesised by a
 CC polypeptide solid-phase synthesis method, and has the monomer sequence of
 CC IEGPTLRQWLAAARAC and the amidated peptide chain structure of
 CC IEGPTLRQWLAAARAC-NH₂. Its activity is increased by 20 times for its
 CC monomer, or by 10 times for the amidated peptide chain compared with the
 CC monomer, or by 100 times for its dimer compared with its monomer
 XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 3; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
 |||||
 Db 1 IEGPTLRQWLAAARA 14

RESULT 30
 AAU25996
 ID AAU25996 standard; peptide; 15 AA.

XX AC AAU25996;

XX DT 17-DEC-2001 (first entry)

XX DE Human thrombopoietin receptor (TPO-R) activator peptide #182.

XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX OS Homo sapiens.

XX PN US6251864-B1.

XX PD 26-JUN-2001.

XX PF 01-MAR-2000; 2000US-00516704.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PR 07-JUN-1996; 96WO-US009623.

XX PR 15-AUG-1996; 96US-00699027.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

XX PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

XX PI Yin Q;

XX DR WPI; 2001-564142/63.

XX PS Disclosure; Col 143-144; 128pp; English.

XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to prevent and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
 |||||
 Db 2 IEGPTLRQWLAAARA 15

RESULT 31
 AAU25831

ID AAU25831 standard; peptide; 15 AA.

XX AC AAU25831;

XX DT 17-DEC-2001 (first entry)

XX DE Human thrombopoietin receptor (TPO-R) activator peptide #17.

XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX OS Homo sapiens.

XX PN US6251864-B1.

XX PD 26-JUN-2001.

XX PF 01-MAR-2000; 2000US-00516704.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PR 07-JUN-1996; 96WO-US009623.

XX PR 15-AUG-1996; 96US-00699027.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

XX PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

XX PI Yin Q;

XX DR WPI; 2001-564142/63.

XX PS Claim 1; Col 69-70; 128pp; English.

XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 CC
 XX Sequence 15 AA;

Query Match 100.0%; Score 73; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 |||||
 Db 1 IEPTLRQWLAARA 14
 |||||

RESULT 32
 ABP51670
 ID ABP51670 standard; peptide; 15 AA.

XX AC ABP51670;

XX DT 01-OCT-2002 (first entry)

XX DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:2.

XX KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 KW complementarity determining region; immunoglobulin; antianaemic;
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.

XX OS Homo sapiens.
 XX OS Synthetic.

XX PN WO200246238-A2.

XX PD 13-JUN-2002.

XX PF 05-DEC-2001; 2001WO-US047656.

XX PR 05-DEC-2000; 2000US-0251448P.

XX PR 04-MAY-2001; 2001US-0288889P.

XX PR 29-MAY-2001; 2001US-0294068P.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX DR WPI; 2002-566610/60.

XX PT A novel immunogen molecule comprising a region in which amino acid
 PT residues corresponding to at least a portion of the complementary
 PT determining region are replaced or fused with an erythropoietin or
 PT thrombopoietin mimetic.

XX PS Claim 19; Page 6; 113pp; English.

XX CC The present invention describes an immunoglobulin molecule or its fragment
 CC (I) comprising a region where amino acid residues corresponding to at
 CC least a portion of the complementary determining region (CDR) are
 CC replaced or fused with biologically active peptides e.g. a peptide
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
 CC that is flanked with proline at its carboxy terminus. (I) has
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as
 CC a stimulator of proliferation, differentiation and maturation of
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
 CC for stimulating proliferation, differentiation or growth of
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with
 CC promegakaryocytes or megakaryocytes, which results in increased platelet

CC production. (I) with a region where amino acid residues corresponding to
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
 CC production of red blood cells, where (I) is contacted with haematopoietic
 CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiency in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
 CC the exemplification of the present invention
 XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 |||||
 Db 1 IEPTLRQWLAARA 14
 |||||

RESULT 33

ABR62908
 ID ABR62908 standard; peptide; 15 AA.

XX AC ABR62908;

XX DT 04-DEC-2003 (first entry)

XX DE Thrombopoietin mimetic peptide AFI2505.

XX KW Thrombopoietin; mimetic; thrombocytopaenia; antibody targeting.

XX OS Synthetic.

XX PN WO2003059251-A2.

XX PD 24-JUL-2003.

XX PF 22-OCT-2002; 2002WO-US033991.

XX PR 22-OCT-2001; 2001US-0344614P.

XX PR 19-SEP-2002; 2002US-0412455P.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Barbas CF, Rader C, Sinha SC, Lerner R;

XX DR WPI; 2003-636673/60.

XX PT Antibody targeting compound useful e.g. for diagnostic immunoassays and
 PT treating microbial diseases comprises targeting or biological agent
 PT covalently linked to combining site of the antibody.

XX PS Example 7; Page 62; 56pp; English.

XX CC The present sequence is that of thrombopoietin (TPO) mimetic peptide
 CC AFI2505, modified to include an N-terminal Cys residue. AFI2505 mimics
 CC the activity of recombinant TPO. The invention provides antibody
 CC targeting compounds that are used to reprogram the specificity of an
 CC antibody. The antibody targeting compound is linked to the combining site
 CC of the antibody, such that the modified antibody takes on the binding
 CC specificity of the targeting agent. In an example from the invention, a
 CC TPO receptor targeting antibody compound was prepared by covalently
 CC linking Cys-modified peptide AFI2505 to aldolase monoclonal antibody 38C2
 CC using a maleimide-diketone linker. The resulting TPO receptor targeting
 CC antibody compound can be used to treat thrombocytopaenia resulting from
 CC chemotherapy and bone marrow transplantation
 XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 7; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
Db 2 IEGPTLRQWLAAARA 15

RESULT 34
ADM72485
ID ADM72485 standard; peptide; 15 AA.
AC ADM72485;
XX
XX
DT 17-JUN-2004 (first entry)
XX
XX DE TPO mimetic peptide fragment.
XX
XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 15
FT /label= bala
FT /note= "beta-alanine"
XX
XX WO2004026332-A1.
XX
XX 01-APR-2004.
XX
XX 18-SEP-2003; 2003WO-US029701.
XX
XX 18-SEP-2002; 2002US-0411700P.
XX 18-SEP-2002; 2002US-0411779P.
XX
XX (THRE-) 3-DIMENSIONAL PHARM INC.
XX
XX Kaushansky K, Macdonald BR;
XX
XX WPI; 2004-283153/26.
XX
XX Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
FT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
XX Disclosure; Fig 2; 32pp; English.

The invention relates to a method (M1) for increasing haematopoietic stem cell production in a subject which involves administering a Thrombopoietin (TPO) mimetic compound to the subject. Also included is another method (M2) of providing haematopoietic stem cells to a subject which involves administering a TPO mimetic compound to a subject to enhance expansion of a stem cell population within bone marrow and/or mobilize stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or the stem cells in the peripheral circulation, and transplanting the harvested stem cells into the subject. A method (M3) is also provided for reducing a time to engraftment following reinfusion of stem cells in a subject, involves administering a TPO mimetic compound to the subject, enhancing the expansion of the stem cell population within bone marrow and/or mobilizing the stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or one or more of the stem cells in the peripheral circulation, and transplanting the one or more harvested stem cells into the subject. TPO mimetic compounds are disclosed as peptides, including cyclic or modified peptides. (M1) is useful for increasing haematopoietic stem cell production in a subject e.g., human. (M3) is useful for reducing time to engraftment following reinfusion of stem cells, reducing the incidence of delayed primary engraftment, reducing the incidence of secondary failure of platelet production and reducing the time of platelet and/or neutrophil engraftment following reinfusion of stem cells in a subject.

CC (M1) is also useful for increasing the number of stem cells from a donor whose cells are then used for rescue of recipient subject. Also useful in the treatment of thrombocytopenia. (M1) enables transplantation to proceed in patients who would not otherwise be considered as candidates because of unacceptably high risk of failed engraftment, reduces the number of aphereses required to generate a minimum acceptable harvest, reduces the incidence of primary and secondary failure of engraftment by increasing the number of haematopoietic stem cells (HSCs) available for transplantation and reduces the time required for primary engraftment. CC The present sequence represents an example of TPO mimetic peptide fragment.
XX
XX
XX Sequence 15 AA;
SQ
Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAARA 14
Db 1 IEGPTLRQWLAAARA 14

RESULT 35
ADM72479
ID ADM72479 standard; peptide; 15 AA.
XX
XX AC ADM72479;
XX
XX 17-JUN-2004 (first entry)
XX
XX TPO mimetic peptide fragment.
XX
XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 15
FT /note= "beta-alanine"
XX
XX WO2004026332-A1.
XX
XX 01-APR-2004.
XX
XX 18-SEP-2003; 2003WO-US029701.
XX
XX 18-SEP-2002; 2002US-0411700P.
XX 18-SEP-2002; 2002US-0411779P.
XX
XX (THRE-) 3-DIMENSIONAL PHARM INC.
XX
XX Kaushansky K, Macdonald BR;
XX
XX WPI; 2004-283153/26.
XX
XX Increasing hematopoietic stem cell production in subject, useful in
PT reducing the incidence of delayed primary engraftment, comprises
FT administering a Thrombopoietin mimetic compound e.g., a peptide to a
PT subject.
XX
XX Disclosure; Fig 2; 32pp; English.

The invention relates to a method (M1) for increasing haematopoietic stem cell production in a subject which involves administering a Thrombopoietin (TPO) mimetic compound to the subject. Also included is another method (M2) of providing haematopoietic stem cells to a subject which involves administering a TPO mimetic compound to a subject to enhance expansion of a stem cell population within bone marrow and/or mobilize stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or the stem cells in the peripheral circulation, and transplanting the harvested stem cells into the subject. A method (M3) is also provided for reducing a time to engraftment following reinfusion of stem cells in a subject, involves administering a TPO mimetic compound to the subject, enhancing the expansion of the stem cell population within bone marrow and/or mobilizing the stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or one or more of the stem cells in the peripheral circulation, and transplanting the one or more harvested stem cells into the subject. TPO mimetic compounds are disclosed as peptides, including cyclic or modified peptides. (M1) is useful for increasing haematopoietic stem cell production in a subject e.g., human. (M3) is useful for reducing time to engraftment following reinfusion of stem cells, reducing the incidence of delayed primary engraftment, reducing the incidence of secondary failure of platelet production and reducing the time of platelet and/or neutrophil engraftment following reinfusion of stem cells in a subject.

CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the incidence of secondary failure
 CC of platelet production and reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.

Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
 Db 1 IEGPTLRQWLAAARA 14

RESULT 36

ADM72478
 ID ADM72478 standard; peptide; 15 AA.

XX AC ADM72478;

XX DT 17-JUN-2004 (first entry)

XX DE TPO mimetic peptide fragment.

XX KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;

XX KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 15
 FT /note= "Lys (15) is linked to one copy of the TPO mimetic
 FT peptide through the alpha amino group and to a second
 FT copy of the peptide (not shown) via the omega amino
 FT group"

FN WO2004026332-A1.

XX PD 01-APR-2004.

XX PF 18-SEP-2003; 2003WO-US029701.

XX PR 18-SEP-2002; 2002US-0411700P.

XX PR 18-SEP-2002; 2002US-0411779P.

XX PA (THRE-) 3-DIMENSIONAL PHARM INC.

XX PI Kaushansky K, Macdonald BR;

XX WPI; 2004-283153/26.

XX Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.

XX Disclosure; Fig 2; 32pp; English.

PS The invention relates to a method (M1) for increasing haematopoietic stem
 CC cell production in a subject which involves administering a
 CC thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing haematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the incidence of secondary failure
 CC of platelet production and reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.

Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14

Db 1 IEGPTLRQWLAAARA 14

RESULT 37

ADM72533
 ID ADM72533 standard; peptide; 15 AA.

XX AC ADM72533;

XX DT 17-JUN-2004 (first entry)

XX DE TPO mimetic peptide fragment.

XX KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;

XX KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Disulfide-bond 1

FT /note = connected to a cysteine residue of a similar TPO
 FT mimetic peptide (not shown) by a disulphide bridge

XX PN WO2004026332-A1.
 XX PD 01-APR-2004.
 XX PF 18-SEP-2003; 2003WO-US029701.
 XX PR 18-SEP-2002; 2002US-0411700P.
 XX PR 18-SEP-2002; 2002US-0411779P.
 XX PA (THRE-) 3-DIMENSIONAL PHARM INC.
 XX PI Kaushansky K, Macdonald BR;
 XX DR WPI; 2004-283153/26.
 XX DR Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX PF Disclosure; Fig 2; 32pp; English.
 XX CC The invention relates to a method (M1) for increasing hematopoietic stem
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing hematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing hematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of hematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 XX fragment.
 XX SQ Sequence 15 AA;
 XX
 XX Query Match 100.08; Score 73; DB 8; Length 15;
 XX Best Local Similarity 100.08; Pred. No. 1.8e-05;
 XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX Qy 1 IEPTLRWLAAARA 14
 XX Db 2 IEPTLRWLAAARA 15
 XX
 XX RESULT 38
 XX ADM72522
 XX ID ADM72522 standard; peptide; 15 AA.
 XX

AC ADM72522;
 XX 17-JUN-2004 (first entry)
 XX DE TPO mimetic peptide fragment.
 XX KW TPO; hematopoietic stem cell; thrombopoietin; haemostatic;
 XX KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FH Modified-site 15
 XX FT /note= "Lys (15) is linked to one copy of the TPO mimetic
 XX FT peptide through the alpha amino group and to a second
 XX FT copy of the peptide (not shown) via the omega amino
 XX FT group"
 XX PN WO2004026332-A1.
 XX PD 01-APR-2004.
 XX PF 18-SEP-2003; 2003WO-US029701.
 XX PR 18-SEP-2002; 2002US-0411700P.
 XX PR 18-SEP-2002; 2002US-0411779P.
 XX PA (THRE-) 3-DIMENSIONAL PHARM INC.
 XX PI Kaushansky K, Macdonald BR;
 XX DR WPI; 2004-283153/26.
 XX DR Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX PF Disclosure; Fig 2; 32pp; English.
 XX CC The invention relates to a method (M1) for increasing hematopoietic stem
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing hematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing hematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of hematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 XX fragment.
 XX SQ Sequence 15 AA;
 XX
 XX Query Match 100.08; Score 73; DB 8; Length 15;
 XX Best Local Similarity 100.08; Pred. No. 1.8e-05;
 XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX Qy 1 IEPTLRWLAAARA 14
 XX Db 2 IEPTLRWLAAARA 15
 XX
 XX RESULT 38
 XX ADM72522
 XX ID ADM72522 standard; peptide; 15 AA.
 XX

CC fragment.
 XX Sequence 15 AA;
 SQ Query Match 100.0%; Score 73; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
 |||||
 DB 1 IEGPTLRQWLAAARA 14
 |||||

RESULT 39
 ADM72523
 ID ADM72523 standard; peptide; 15 AA.
 XX
 AC ADM72523;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE TPO mimetic peptide fragment.
 XX
 XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
 XX
 OS Synthetic.

Key Location/Qualifiers
 FH Modified-site 15
 FT /note= "beta-alanine"
 FT
 XX WO2004026332-A1.
 PN
 XX 01-APR-2004.
 PD
 XX 18-SEP-2003; 2003WO-US029701.
 PF
 XX 18-SEP-2002; 2002US-0411700P.
 PR 18-SEP-2002; 2002US-0411779P.
 XX (THRE-) 3-DIMENSIONAL PHARM INC.
 PA Kaushansky K, Macdonald BR;
 XX WPI; 2004-283153/26.
 XX Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX Disclosure; Fig 2; 32pp; English.

CC The invention relates to a method (M1) for increasing haematopoietic stem
 CC cell production in a subject which involves administering a
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing haematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to

CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the incidence of secondary failure
 CC of platelet production and reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells from a donor
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 73; DB 8; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAARA 14
 |||||
 DB 1 IEGPTLRQWLAAARA 14
 |||||

RESULT 40
 ADM72482
 ID ADM72482 standard; peptide; 15 AA.
 XX
 AC ADM72482;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE TPO mimetic peptide fragment.
 XX
 XX TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
 XX
 OS Synthetic.

Key Location/Qualifiers
 FH Modified-site 15
 FT /note= "Lys (15) is linked to one copy of the TPO mimetic
 FT peptide through the alpha amino group and to a second
 FT copy of the peptide (not shown) via the omega amino
 FT group"
 XX WO2004026332-A1.
 PN
 XX 01-APR-2004.
 PD
 XX 18-SEP-2003; 2003WO-US029701.
 PF
 XX 18-SEP-2002; 2002US-0411700P.
 PR 18-SEP-2002; 2002US-0411779P.
 XX (THRE-) 3-DIMENSIONAL PHARM INC.
 PA Kaushansky K, Macdonald BR;
 XX WPI; 2004-283153/26.
 XX Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX Disclosure; Fig 2; 32pp; English.
 XX The invention relates to a method (M1) for increasing haematopoietic stem
 XX cell production in a subject which involves administering a

CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
CC another method (M2) of providing haematopoietic stem cells to a subject
CC which involves administering a TPO mimetic compound to a subject to
CC enhance expansion of a stem cell population within bone marrow and/or
CC mobilize stem cells in peripheral circulation, harvesting one or more of
CC the bone marrow stem cells or the stem cells in the peripheral
CC circulation, and transplanting the harvested stem cells into the subject.
CC A method (M3) is also provided for reducing a time to engraftment
CC following reinfusion of stem cells in a subject, involves administering a
CC TPO mimetic compound to the subject, enhancing the expansion of the stem
CC cell population within bone marrow and/or mobilizing the stem cells in
CC peripheral circulation, harvesting one or more of the bone marrow stem
CC cells or one or more of the stem cells in the peripheral circulation, and
CC transplanting the one or more harvested stem cells into the subject. TPO
CC mimetic compounds are disclosed as peptides, including cyclic or modified
CC peptides. (M1) is useful for increasing haematopoietic stem cell
CC production in a subject e.g., human. (M3) is useful for reducing time to
CC engraftment following reinfusion of stem cells, reducing the incidence of
CC delayed primary engraftment, reducing the incidence of secondary failure
CC of platelet production and reducing the time of platelet and/or
CC neutrophil engraftment following reinfusion of stem cells in a subject.
CC (M1) is also useful for increasing the number of stem cells from a donor
CC whose cells are then used for rescue of recipient subject. Also useful in
CC the treatment of thrombocytopenia. (M1) enables transplantation to
CC proceed in patients who would not otherwise be considered as candidates
CC because of unacceptably high risk of failed engraftment, reduces the
CC number of aphereses required to generate a minimum acceptable harvest,
CC reduces the incidence of primary and secondary failure of engraftment, by
CC increasing the number of haematopoietic stem cells (HSCs) available for
CC transplantation and reduces the time required for primary engraftment.
CC The present sequence represents an example of TPO mimetic peptide
CC fragment.

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 41
ADQ16585
ID ADQ16585 standard; peptide; 15 AA.

XX AC ADQ16585;

XX DT 09-SEP-2004 (first entry)

XX DE TPO mimetic peptide SEQ ID NO:2.

XX KW immunoglobulin; complementarity determining region; CDR; peptide mimetic;
XX KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
XX KW immunotherapy; thrombocytopenia.

XX OS Unidentified.

XX PN WO2004050017-A2.

XX PD 17-JUN-2004.

XX PF 17-NOV-2003; 2003WO-US036894.

XX PR 02-DEC-2002; 2002US-00307724.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Frederickson S, Renshaw M;

XX DR WPI; 2004-460973/43.

XX New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
or a TPO mimetic, useful for treating thrombocytopenia.

XX Disclosure; SEQ ID NO 2; 107pp; English.

XX The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced
CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
CC invention has immunosuppressive activity, and may have a use in
CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
CC treating thrombocytopenia as a result of chemotherapy, bone marrow
CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
CC The present sequence represents a TPO mimetic peptide.

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 42

ADT92483

ID ADT92483 standard; peptide; 15 AA.

XX AC ADT92483;

XX DT 13-JAN-2005 (first entry)

XX DE Modified thrombopoietin mimetic peptide AFI2505 with amino Cys residue.

XX KW integrin targeting compound; RGD peptidomimetic; integrin-alpha;
XX KW angiogenesis; bone metabolism; inflammation; cancer; cell growth;
XX KW antiangiogenic; osteopathic; antiinflammatory; cytostatic; TPO;
XX KW thrombopoietin.

XX OS Synthetic.

XX OS Unidentified.

XX PN WO2004091542-A2.

XX PD 28-OCT-2004.

XX PF 15-APR-2004; 2004WO-US012034.

XX PR 15-APR-2003; 2003US-0463456P.

XX PR 30-SEP-2003; 2003US-0507887P.

XX PA (COVX-) COVX PHARM INC.

XX PI Tamiz A, Bradshaw CW;

XX WPI; 2004-784488/77.

XX New nitrogen containing integrin targeting compounds are integrin alpha-v
PT beta-3 antagonists useful for the treatment or prevention of e.g. bone
PT metabolism, inflammation, cancer and cell growth.

XX Example 7; SEQ ID NO 6; 182pp; English.

XX This invention relates to novel nitrogen containing integrin targeting
CC compounds, their stereoisomers, tautomers or salts thereof. Specifically,
CC it refers to integrin targeting compounds that comprise small MW, RGD
CC peptidomimetic integrin targeting agent-linker conjugates that are linked
CC to polymers such as a protein. The present invention describes compounds

CC that can be useful for the treatment or prevention of diseases that
CC involve integrin-alpha, and can be used to treat a defect in
CC angiogenesis, bone metabolism, inflammation, cancer or cell growth.
CC Accordingly, they exhibit antiangiogenic, osteopathic, antiinflammatory
CC and cytostatic activities. This peptide sequence is a modified TPO
CC mimetic peptide with amino Cys given in an exemplification of the
CC invention.

XX Sequence 15 AA;
SQ

Query Match 100.0%; Score 73; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 2 IEGPTLRQWLAARA 15

RESULT 43
ADU70209
ID ADU70209 standard; peptide; 15 AA.
XX
AC ADU70209;
XX
DT 10-FEB-2005 (first entry)
XX
DE Thrombopoietin receptor modulator peptide #4.
XX
KW hematological disease; drug delivery; thrombopoietin receptor modulator;
KW diagnosis; pharmaceutical.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Ile is acetylated"
FT Modified-site 15 /label= OTHER
FT /note= "Covalently linked to a spacer/PEG moiety which is
FT in turn lionked to another copy of this sequence forming
FT a dimer"
XX
PN WO2004100997-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014887.
XX
PR 12-MAY-2003; 2003US-0469996P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Holmes C, Yin K, Tumelty D;
XX
WPI; 2005-039305/04.
XX
PT New peptide-based compound comprising a peptide moiety, a spacer moiety
PT and a water-soluble polymer moiety useful as pharmaceuticals.
XX
PS Example 14; Page 43; 57pp; English.
XX
CC The invention relates to a peptide-based compound (I) comprising a
CC peptide moiety, a spacer moiety and a water-soluble polymer moiety (e.g.
CC polyethylene glycol, PEG). Also included is a pharmaceutical composition
CC comprising (I) and one or more pharmaceutically acceptable diluents,
CC preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. The
CC peptide moiety comprises one or more peptides that bind to erythropoietin
CC -receptors or thrombopoietin-receptors. The peptide based compound is
CC useful as a pharmaceutical. The peptide moiety particularly
CC erythropoietin receptors are essential in the process of red blood cell
CC formation and has useful applications in both diagnosis and the treatment
CC of blood disorders characterized by low or defective red blood cell

CC production. The peptide based compound has improved antigenicity and
CC durability. Polyethylene glycolylation leads to improved drug efficacy and
CC safety, increases bioavailability and decreases immunogenicity and dosing
CC frequency. The present sequence is a thrombopoietin-receptor modulating
CC peptide.

XX Sequence 15 AA;
SQ

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 44
ADU75981
ID ADU75981 standard; peptide; 15 AA.
XX
AC ADU75981;
XX
DT 10-FEB-2005 (first entry)
XX
DE Peptide-based compound associated peptide #2.
XX
KW antianemic; pharmaceutical; blood; blood disorder.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "N-terminal acetyl"
FT Modified-site 15 /label= OTHER
FT /note= "OTHER= Two of these molecules are optionally
FT attached by the C-terminal by a spacer molecule that can
FT undergo PEGylation"
XX
PN WO2004101600-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014888.
XX
PR 12-MAY-2003; 2003US-0470246P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Yin K, Holmes C, Lalonde G, Balu P, Tumelty D, Schatz P;
XX
WPI; 2005-039327/04.
XX
PT Peptide-based compound useful as e.g. pharmaceutical for the treatment of
PT blood disorders comprises a peptide moiety and a linear
PT poly(ethyleneglycol) moiety.
XX
PS Disclosure; Page 44; 56pp; English.
XX
CC The invention describes a peptide-based compound (A) comprising a peptide
CC moiety (B) and a linear poly(ethyleneglycol) moiety (C) having molecular
CC weight of more than 20 KDaltons. The compounds are useful as
CC pharmaceuticals for the treatment of blood disorders characterized by low
CC or defective red blood cell production, for reducing immunogenicity and
CC as peptide modifying reagent. The compounds exhibit improved antigenicity
CC and durability. This is the amino acid sequence of peptide that may be
CC present in the peptide moieties of the invention.

XX Sequence 15 AA;
SQ

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLARA 14

RESULT 45
ADV44320
ID ADV44320 standard; peptide; 15 AA.
XX
AC ADV44320;
XX
DT 10-MAR-2005 (first entry)
XX
DE Agonist TPO mimetic peptide seqid 2.
XX
KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
PN WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orenicia C;
XX
XX WPI; 2005-031588/03.
XX
DR New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Disclosure; SEQ ID NO 2; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least
CC a portion of two CDRs are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to
CC replace a CDR in a rationally designed antibody.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLARA 14

RESULT 46
AEB12793
ID AEB12793 standard; peptide; 15 AA.
XX
AC AEB12793;
XX
DT 08-SEP-2005 (first entry)
XX
DE TPO mimetic peptide, SEQ ID 2.
XX
KW TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.
XX
OS Synthetic.
XX
PN WO2005060642-A2.
XX
PD 07-JUL-2005.
XX
PF 15-DEC-2004; 2004WO-US041946.
XX
PR 15-DEC-2003; 2003US-00737290.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orenicia C;
XX
XX WPI; 2005-479402/48.
XX
DR New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.
XX
PS Disclosure; SEQ ID NO 2; 152pp; English.
XX
CC The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, extendin, extendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under
CC conditions suitable for expression of the immunoglobulin or its
CC fragment), a composition comprising an immunoglobulin (or its fragment)
CC and a pharmaceutically acceptable carrier, treating congestive heart
CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, extendin, extendin mimetics,
CC glucagons, glucagon mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1),
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to

CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a thrombopoietin (TPO) mimetic, erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.

XX SQ Sequence 15 AA;
Query Match 100.0%; Score 73; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 47
AAW19534
ID AAW19534 standard; protein; 16 AA.

XX AC AAW19534;

XX DT 10-SEP-1997 (first entry)

XX DE Thrombopoietin receptor binding compound peptide (part of a dimer).

XX KW Haematology; thrombocytopenia; TPO; TR; proliferation;
XX bone marrow transfusion; chemotherapy; radiation therapy.

XX OS Synthetic.

XX FH Key Location/Qualifiers
FT Modified-site 15
FT Cross-links /label= bAla 16
FT Modified-site 16 /note= "Linked to the omega Ala in AAW09468"
FT /note= "In amide form"

XX PN WO9640189-Al.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX PA Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CW, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.

PS Claim 30; Page 91; 106pp; English.

XX CC The present sequence is a compound which binds to thrombopoietin (TPO)
XX receptor (TR). It is part of a dimer linked by the omega amino acid to
XX the omega amino acid in the sequence in AAW09468. The compound can be
XX used for treating patients suffering from haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. The peptide may also be used to maintain the
XX proliferation and growth of TPO-dependent cell lines and for use in
XX biological research, for detecting TPO receptors on living cells

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAARA 14
| | | | | | | | | |
Db 1 IEGPTLRQWLAARA 14

RESULT 48
AAW33035
ID AAW33035 standard; peptide; 16 AA.

XX AC AAW33035;

XX DT 11-MAR-1998 (first entry)

XX DE Thrombopoietin receptor binding peptide.

XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;

XX KW haematological disorder; thrombocytopenia; chemotherapy;

XX KW radiation therapy; bone marrow transfusion; diagnosis;

XX KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX FH Key Location/Qualifiers
FT Cross-links 14
FT /note= "epsilon amino group of Lys16 linked to terminal
FT carboxy group of AAW33034"

FT Modified-site 15
FT /label= bAla

XX PN WO9640750-Al.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CW, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 30; Page 91; 106pp; English.

XX CC The present peptide binds the thrombopoietin receptor (TR), has a
XX molecular weight of less than 8000 Da and a TR binding affinity as
XX expressed by an IC50 of no more than about 100 microm. It can be used to
XX treat disorders which are susceptible to treatment with a thrombopoietin
XX agonist, preferably haematological disorders and thrombocytopenia

CC resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. It can also be used diagnostically, e.g. to investigate the
 CC mechanism of thrombopoietin signal transduction and receptor activation,
 CC or to maintain the proliferation and growth of thrombopoietin dependent
 CC cell lines

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 |||||
 Db 1 IEPTLRQWLAARA 14

RESULT 49

AAW36775
 ID AAW36775 standard; peptide; 16 AA.

XX AC AAW36775;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 XX haematological disorder; thrombocytopaenia; chemotherapy;
 XX radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX Key Location/Qualifiers
 FH Disulfide-bond 1..16
 FT Modified-site 16
 FT /note= "NH2-Cys"

XX WO9640750-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 77; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 |||||
 Db 2 IEPTLRQWLAARA 15

RESULT 50

AAW36771
 ID AAW36771 standard; peptide; 16 AA.

XX AC AAW36771;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX Key Location/Qualifiers
 FH Disulfide-bond 1..16
 FT Modified-site 16
 FT /note= "NH2-Cys"

XX WO9640750-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 76; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines

SQ Sequence 16 AA;

Query Match 100.0%; Score 73; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 |||||
 Db 2 IEPTLRQWLAARA 15

Search completed: May 12, 2006, 10:36:51

Job time : 105.809 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16.9915 Seconds

(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-13

Perfect score: 73

Sequence: 1 IEPTLRQWLARA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : PIR 80.*

1: Pirl.*

2: Pirl.*

3: Pirl.*

4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.5	67.8	333	2 A36925	transcription acti
2	47	64.4	296	2 AG0147	probable membrane
3	46	63.0	306	2 D70601	UTP-glucose-1-phos
4	44	60.3	200	2 T23485	hypothetical prote
5	44	60.3	207	2 T37464	probable glutathio
6	44	60.3	536	1 SYECBB	2,3-dihydroxybenzo
7	44	60.3	536	2 E85558	2,3-dihydroxybenzo
8	44	60.3	536	2 A99708	probable dimethyla
9	43	58.9	285	2 G71337	conserved hypothet
10	43	58.9	683	2 B71325	hypothetical prote
11	42	57.5	473	2 E84853	pol polyprotein -
12	42	57.5	1019	2 T11560	probable phosphopa
13	41	56.2	195	2 F91171	probable phosphopa
14	41	56.2	195	2 F86017	hypothetical 21.8K
15	41	56.2	195	2 S47694	ABC transporter, A
16	41	56.2	249	2 E87575	UTP-glucose-1-phos
17	41	56.2	306	2 T45453	glyceraldhyde-3-p
18	41	56.2	326	2 C24430	glyceraldhyde-3-p
19	41	56.2	337	1 DEP2G	glyceraldhyde-3-p
20	41	56.2	337	2 A35080	glyceraldhyde-3-p
21	41	56.2	338	1 DEIS3C	glyceraldhyde-3-p
22	41	56.2	338	2 JQ1287	conserved hypothet
23	41	56.2	719	2 B95325	topoisomerase iv c
24	41	56.2	750	2 A97501	topoisomerase IV s
25	41	56.2	750	2 AE2719	simx protein - str
26	40	54.8	239	2 S25204	hypothetical prote
27	40	54.8	463	2 S27491	hypothetical prote
28	40	54.8	530	2 A81958	probable permease
29	40	54.8	531	2 E81015	ABC transporter, p

30	40	54.8	656	2	S30484	pol polyprotein -
31	40	54.8	656	2	S30483	pol polyprotein -
32	40	54.8	721	2	A39707	erythrocyte membra
33	40	54.8	1123	2	T51517	telomerase reverse
34	40	54.8	1712	1	CGHU2B	collagen alpha 2(I
35	39.5	54.1	325	2	A84326	hypothetical prote
36	39	53.4	131	2	S74539	hypothetical prote
37	39	53.4	217	2	S46354	pol polyprotein -
38	39	53.4	267	2	I40327	baf protein - Bord
39	39	53.4	331	2	B48445	glyceraldhyde-3-p
40	39	53.4	331	2	A72514	hypothetical prote
41	39	53.4	400	2	C87021	serine-threonine p
42	39	53.4	600	2	C83221	transport protein
43	39	53.4	791	2	A82291	c-di-GMP phosphodi
44	39	53.4	1034	1	GNLJCA	HIV-1 retropepsin
45	39	53.4	1035	1	GNLJGG	HIV-1 retropepsin
46	39	53.4	1036	1	GNLJG2	HIV-1 retropepsin
47	39	53.4	1055	1	GNLJST	HIV-1 retropepsin
48	39	53.4	1058	2	S08436	pol polyprotein -
49	39	53.4	1058	2	S08436	hypothetical prote
50	39	53.4	3345	2	T13423	transcription regu
51	38	52.1	134	2	B75468	probable oxidoredu
52	38	52.1	197	2	G82973	glyceraldhyde-3-p
53	38	52.1	246	2	AH0190	glyceraldhyde-3-p
54	38	52.1	247	2	FQ0178	integrase/recombin
55	38	52.1	295	2	T07730	regulatory protein
56	38	52.1	297	2	B87109	transcription acti
57	38	52.1	311	1	RGECK	positive regulator
58	38	52.1	311	2	AH0867	positive regulator
59	38	52.1	311	2	C85936	hypothetical prote
60	38	52.1	311	2	H91090	hypothetical prote
61	38	52.1	314	2	H70723	glyceraldhyde-3-p
62	38	52.1	337	1	DEP2G	glyceraldhyde-3-p
63	38	52.1	337	1	DESKG	glyceraldhyde-3-p
64	38	52.1	337	1	DEUSGM	glyceraldhyde-3-p
65	38	52.1	337	1	DEZMGC	hypothetical prote
66	38	52.1	339	2	A83358	hypothetical prote
67	38	52.1	350	2	B38535	A/G-specific adeni
68	38	52.1	350	2	H85953	adenine glycosylas
69	38	52.1	350	2	E91108	adenine glycosylas
70	38	52.1	360	2	S38570	hypothetical prote
71	38	52.1	469	2	AD1926	hypothetical prote
72	38	52.1	589	2	F87626	sensor histidine k
73	38	52.1	635	2	A87433	hypothetical prote
74	38	52.1	816	2	A71006	hypothetical prote
75	38	52.1	904	2	C70559	probable polA prot
76	38	52.1	1155	2	AC2426	adenylate cyclase
77	37.5	51.4	436	2	JC4742	transposase - Cory
78	37	50.7	151	2	S63748	Hiv-1 retropepsin
79	37	50.7	165	2	F87542	hypothetical prote
80	37	50.7	234	2	P00179	glyceraldhyde-3-p
81	37	50.7	305	2	A24159	glyceraldhyde-3-p
82	37	50.7	335	2	S29813	glyceraldhyde-3-p
83	37	50.7	337	1	DEBHG	glyceraldhyde-3-p
84	37	50.7	337	2	S42479	glyceraldhyde-3-p
85	37	50.7	337	2	T02723	glyceraldhyde-3-p
86	37	50.7	337	2	T02722	glyceraldhyde-3-p
87	37	50.7	338	1	DENDG	glyceraldhyde-3-p
88	37	50.7	341	1	DEJMG	glyceraldhyde-3-p
89	37	50.7	341	2	T08147	glyceraldhyde-3-p
90	37	50.7	341	2	AG0195	probable exported
91	37	50.7	341	2	AG0195	conserved hypothet
92	37	50.7	352	2	G83636	hypothetical prote
93	37	50.7	391	2	T36739	hypothetical prote
94	37	50.7	407	2	A86298	hypothetical prote
95	37	50.7	422	2	F96826	glyceraldhyde-3-p
96	37	50.7	433	2	S51837	glyceraldhyde-3-p
97	37	50.7	438	2	G87337	membrane protein,
98	37	50.7	480	2	H84747	probable steroid d
99	37	50.7	486	2	B86411	protein F3M18.4 [i
100	37	50.7	544	2	A72459	probable glutamyl-
101	37	50.7	547	2	T36550	hypothetical prote
102	37	50.7	547	2	T36550	hypothetical prote

103	37	50.7	664	2	H83962	serine/threonine p	176	36	49.3	2825	2	T14271	Doc4 protein, stre
104	37	50.7	762	2	H83415	cis/trans isomeras	177	35.5	48.6	602	2	T45278	oligopeptide ABC t
105	37	50.7	777	2	A87309	hypothetical prote	178	35	47.9	70	2	T06920	ribosomal protein
106	37	50.7	825	2	JC4163	DNA-binding protei	179	35	47.9	120	2	A97655	conserved hypotet
107	37	50.7	921	2	A40597	pol polyprotein -	180	35	47.9	120	2	AG2878	hypothetical prote
108	37	50.7	1009	2	S28081	pol polyprotein -	181	35	47.9	125	2	C83138	NADH2 dehydrogenas
109	37	50.7	1039	2	S46347	regulator protein	182	35	47.9	129	2	S28237	probable insertion
110	37	50.7	1083	2	S54293	major core protein	183	35	47.9	150	2	T34697	hypothetical prote
111	37	50.7	1233	1	F3XR03	probable polyketid	184	35	47.9	198	2	B84382	hypothetical prote
112	37	50.7	2108	2	H70819	CSOC3.6 protein -	185	35	47.9	209	2	D95274	hypothetical prote
113	37	50.7	2329	2	S44625	probable pre-mRNA	186	35	47.9	211	2	AB3504	hypothetical prote
114	37	50.7	2363	2	T38841	probable UV damage	187	35	47.9	216	2	D75567	GTP cyclohydrolase
115	36.5	50.0	326	2	C75350	tryptophan halogen	188	35	47.9	232	2	B82987	probable hydrolase
116	36.5	50.0	502	2	E87596	ribosomal protein	189	35	47.9	233	2	A83862	initiation of chro
117	36	49.3	60	2	A48396	MLCL536.14 protein	190	35	47.9	256	2	G82549	phage-related prot
118	36	49.3	102	2	T11006	hypothetical prote	191	35	47.9	260	2	S72748	Bl17_FJ_136 prote
119	36	49.3	104	2	A96002	NADH2 dehydrogenas	192	35	47.9	298	2	A70926	glycine cleavage s
120	36	49.3	129	2	J80383	hypothetical prote	193	35	47.9	305	2	AD0126	nosX1 protein [imp
121	36	49.3	153	2	A96751	hypothetical prote	194	35	47.9	323	2	AF3630	probable transposo
122	36	49.3	188	2	C82863	recombinase XFa001	195	35	47.9	327	2	E82277	conserved hypotet
123	36	49.3	189	2	F95017	hypothetical prote	196	35	47.9	333	2	G83391	protein export [im
124	36	49.3	189	2	G97890	hypothetical prote	197	35	47.9	344	2	E84377	hypothetical prote
125	36	49.3	204	2	T44257	thiamine-phosphate	198	35	47.9	355	2	T20825	protein serine/thr
126	36	49.3	207	2	B75327	hypothetical prote	199	35	47.9	362	2	A64212	probable transport
127	36	49.3	220	2	AC0318	probable nicotinat	200	35	47.9	377	2	F95982	probable nuclear D
128	36	49.3	221	2	F81876	hypothetical prote	201	35	47.9	392	2	T51151	probable malate de
129	36	49.3	235	2	C83842	hypothetical prote	202	35	47.9	400	2	T18570	lambda integrase f
130	36	49.3	278	2	T45494	probable transposo	203	35	47.9	415	2	I39527	aminotransferase,
131	36	49.3	302	2	T36030	probable fructokin	204	35	47.9	425	2	A87517	conserved hypotet
132	36	49.3	312	2	F86876	hypothetical prote	205	35	47.9	433	2	C75354	HYPOPHETICAL 50.8
133	36	49.3	327	2	H82736	hypothetical prote	206	35	47.9	451	2	A95379	deoxyribodipyrimid
134	36	49.3	361	2	F87286	cation efflux fami	207	35	47.9	472	1	WZEC0	hypothetical prote
135	36	49.3	370	2	AB2578	membrane-bound lyt	208	35	47.9	472	2	C85571	hypothetical prote
136	36	49.3	370	2	A97360	outer membrane lip	209	35	47.9	472	2	E90720	deoxyribodipyrimid
137	36	49.3	373	2	B95291	conserved hypotet	210	35	47.9	473	2	A10587	deoxyribodipyrimid
138	36	49.3	408	2	D75252	conserved hypotet	211	35	47.9	473	2	S22321	probable transport
139	36	49.3	410	1	DEPSA	3-methyl-2-oxobuta	212	35	47.9	475	2	F84261	probable MFS trans
140	36	49.3	410	1	C83365	2-oxoisovalerate d	213	35	47.9	477	2	H83588	deoxyribodipyrimid
141	36	49.3	458	2	B83372	hypothetical prote	214	35	47.9	501	2	S48120	deoxyribodipyrimid
142	36	49.3	472	2	T20454	hypothetical prote	215	35	47.9	505	2	T19971	hypothetical prote
143	36	49.3	475	2	G70861	hypothetical prote	216	35	47.9	506	2	T19973	hypothetical prote
144	36	49.3	478	2	E90930	hypothetical prote	217	35	47.9	518	2	AD2315	hypothetical prote
145	36	49.3	478	2	A85779	hypothetical prote	218	35	47.9	521	2	T01923	hypothetical prote
146	36	49.3	478	2	B64929	hypothetical prote	219	35	47.9	524	2	E71881	hypothetical prote
147	36	49.3	480	2	AP0704	conserved hypotet	220	35	47.9	527	2	B64633	hypothetical prote
148	36	49.3	481	2	T49321	related to heterok	221	35	47.9	571	2	AH8737	probable histone a
149	36	49.3	491	2	C98275	polyketide synthas	222	35	47.9	578	2	AI1020	cytochrome c-type
150	36	49.3	491	2	AC3009	polyketide synthas	223	35	47.9	605	2	AG0460	probable carbon st
151	36	49.3	564	2	T37934	conserved hypotet	224	35	47.9	610	2	A86176	hypothetical prote
152	36	49.3	586	2	B84271	glutamyl-tRNA synt	225	35	47.9	612	2	S71221	probable deoxyribo
153	36	49.3	664	2	G89894	protein kinase [im	226	35	47.9	628	2	S35547	double-strand-brea
154	36	49.3	691	2	A54741	erythrocyte membra	227	35	47.9	631	2	A54659	DNA repair protein
155	36	49.3	709	2	S75212	comE protein - Syn	228	35	47.9	633	2	A36353	DNA repair protein
156	36	49.3	827	2	S29955	surface glycoprote	229	35	47.9	645	2	G88130	protein F10G7.4 [i
157	36	49.3	855	1	JQ2003	env polyprotein -	230	35	47.9	722	2	T37970	probable G2-specif
158	36	49.3	855	1	JQ2004	env polyprotein -	231	35	47.9	793	2	JC5831	kinesin-related pr
159	36	49.3	855	1	F45557	external glycoprot	232	35	47.9	795	2	H70934	hypothetical prote
160	36	49.3	856	1	A45394	env polyprotein pr	233	35	47.9	803	2	B82392	phosphoenolpyruvat
161	36	49.3	856	1	VCLJ39	env polyprotein pr	234	35	47.9	994	2	AC3480	DNA-directed DNA p
162	36	49.3	856	2	S23823	env polyprotein -	235	35	47.9	1040	2	T08190	hypothetical prote
163	36	49.3	857	2	S19886	env polyprotein E	236	35	47.9	1091	2	G59435	hypothetical prote
164	36	49.3	907	2	AB1885	hypothetical prote	237	35	47.9	1380	2	T25284	hypothetical prote
165	36	49.3	984	2	T48216	hypothetical prote	238	35	47.9	2529	2	B64635	toxin-like outer m
166	36	49.3	1032	2	S12153	pol polyprotein -	239	35	47.9	3172	2	S22012	erythronolide synt
167	36	49.3	1056	1	GNLJG3	HIV-1 retropepsin	240	35	47.9	3178	2	S13595	6-deoxyerythronoli
168	36	49.3	1099	1	S31926	myosin IB heavy ch	241	34.5	47.3	193	2	D83417	probable transcrip
169	36	49.3	1099	2	A59300	myosin-II - mouse	242	34.5	47.3	235	2	S55883	CCHH finger protei
170	36	49.3	1191	2	T31031	hypothetical prote	243	34.5	47.3	382	2	T34789	hypothetical prote
171	36	49.3	1257	2	S28764	neurocan precursor	244	34.5	47.3	409	2	T02776	conserved hypotet
172	36	49.3	1339	2	T40245	probable transcrip	245	34.5	47.3	478	2	AG0170	sarcosine dehydrog
173	36	49.3	1576	2	T28655	receptor-like hist	246	34.5	47.3	774	2	AF3337	isoleucyl-tRNA syn
174	36	49.3	1576	2	T30898	receptor-like hist	247	34.5	47.3	943	2	C82559	hypothetical prote
175	36	49.3	2617	2	AB2136	peptide synthetase	248	34	46.6	85	2	F75291	

249	34	46.6	98	2	A70301	ribosomal protein	322	34	46.6	219	2	S32056	RNA-directed DNA p
250	34	46.6	114	2	A10294	probable membrane	323	34	46.6	219	2	S32059	RNA-directed DNA p
251	34	46.6	119	2	S75551	hypothetical prote	324	34	46.6	219	2	S32131	RNA-directed DNA p
252	34	46.6	129	2	S63743	HIV-1 retropepsin	325	34	46.6	219	2	S32071	RNA-directed DNA p
253	34	46.6	130	2	S63754	HIV-1 retropepsin	326	34	46.6	219	2	S32127	RNA-directed DNA p
254	34	46.6	132	2	S63733	HIV-1 retropepsin	327	34	46.6	219	2	S32134	RNA-directed DNA p
255	34	46.6	136	2	S63746	HIV-1 retropepsin	328	34	46.6	219	2	S32093	RNA-directed DNA p
256	34	46.6	140	2	S63737	HIV-1 retropepsin	329	34	46.6	219	2	S32075	RNA-directed DNA p
257	34	46.6	141	2	S63750	HIV-1 retropepsin	330	34	46.6	219	2	S32079	RNA-directed DNA p
258	34	46.6	142	2	AC0928	probable regulator	331	34	46.6	219	2	S32119	RNA-directed DNA p
259	34	46.6	143	2	S63735	HIV-1 retropepsin	332	34	46.6	219	2	S32138	RNA-directed DNA p
260	34	46.6	145	2	S63745	HIV-1 retropepsin	333	34	46.6	219	2	S32129	RNA-directed DNA p
261	34	46.6	145	2	S07957	hypothetical prote	334	34	46.6	219	2	S32097	RNA-directed DNA p
262	34	46.6	149	2	S63734	HIV-1 retropepsin	335	34	46.6	219	2	S32117	RNA-directed DNA p
263	34	46.6	151	2	S63731	HIV-1 retropepsin	336	34	46.6	219	2	S32077	RNA-directed DNA p
264	34	46.6	153	2	A97524	hypothetical prote	337	34	46.6	219	2	S32098	RNA-directed DNA p
265	34	46.6	158	2	A10605	conserved hypothet	338	34	46.6	219	2	S32051	RNA-directed DNA p
266	34	46.6	158	2	E64823	hypothetical prote	339	34	46.6	219	2	S32132	RNA-directed DNA p
267	34	46.6	158	2	E30745	probable sensory t	340	34	46.6	219	2	S32058	RNA-directed DNA p
268	34	46.6	160	2	S63753	HIV-1 retropepsin	341	34	46.6	219	2	S32140	RNA-directed DNA p
269	34	46.6	162	2	S63732	HIV-1 retropepsin	342	34	46.6	223	2	S44974	lmbU protein - Str
270	34	46.6	174	2	UCL153	hypothetical 19.7K	343	34	46.6	244	2	D86275	F7A19.26 protein -
271	34	46.6	176	2	S63747	HIV-1 retropepsin	344	34	46.6	254	2	F87652	hypothetical prote
272	34	46.6	178	2	AB2743	hypothetical prote	345	34	46.6	254	2	C95273	hypothetical prote
273	34	46.6	180	2	T09063	hypothetical prote	346	34	46.6	255	2	A45861	MHC class II histo
274	34	46.6	184	2	H83409	hypothetical prote	347	34	46.6	259	2	AB3572	succinoglycan bios
275	34	46.6	193	2	A33835	SEG63 protein comp	348	34	46.6	274	2	T05509	hypothetical prote
276	34	46.6	197	2	AS2063	hypothetical prote	349	34	46.6	277	2	I51117	proopiomelanocorti
277	34	46.6	206	2	S38626	glutathione transf	350	34	46.6	281	2	H97548	(ac007190) f23n19
278	34	46.6	215	2	D75379	hypothetical prote	351	34	46.6	281	2	AF2768	phytoene synthase
279	34	46.6	219	2	S32096	RNA-directed DNA p	352	34	46.6	289	2	JQ0059	hypothetical prote
280	34	46.6	219	2	S32157	RNA-directed DNA p	353	34	46.6	295	2	S76136	class I histocompa
281	34	46.6	219	2	S32080	RNA-directed DNA p	354	34	46.6	302	2	E75414	peCM-related prote
282	34	46.6	219	2	S32062	RNA-directed DNA p	355	34	46.6	308	2	AC0098	transcription acti
283	34	46.6	219	2	S32047	RNA-directed DNA p	356	34	46.6	313	2	AC0098	probable LysR-fam
284	34	46.6	219	2	S32065	RNA-directed DNA p	357	34	46.6	315	2	G95415	probable transcrip
285	34	46.6	219	2	S32122	RNA-directed DNA p	358	34	46.6	320	2	T35265	probable D-amino a
286	34	46.6	219	2	S32060	RNA-directed DNA p	359	34	46.6	321	2	C70653	probable prephenat
287	34	46.6	219	2	S32118	RNA-directed DNA p	360	34	46.6	322	2	F86918	probable prephenat
288	34	46.6	219	2	S32057	RNA-directed DNA p	361	34	46.6	324	2	F95306	probable glutamate
289	34	46.6	219	2	S32048	RNA-directed DNA p	362	34	46.6	331	2	F82131	glyceralddehyde 3-p
290	34	46.6	219	2	S32094	RNA-directed DNA p	363	34	46.6	332	2	T44437	aminooxycholester
291	34	46.6	219	2	S32061	RNA-directed DNA p	364	34	46.6	333	1	DELOG3	glyceralddehyde-3-p
292	34	46.6	219	2	S32084	RNA-directed DNA p	365	34	46.6	336	1	DEASG3	glyceralddehyde-3-p
293	34	46.6	219	2	S32152	RNA-directed DNA p	366	34	46.6	337	1	DSJGJC	glyceralddehyde-3-p
294	34	46.6	219	2	S32089	RNA-directed DNA p	367	34	46.6	337	1	DSVDCG	glyceralddehyde-3-p
295	34	46.6	219	2	S32073	RNA-directed DNA p	368	34	46.6	337	2	S26946	glyceralddehyde-3-p
296	34	46.6	219	2	S32049	RNA-directed DNA p	369	34	46.6	337	2	S29814	glyceralddehyde-3-p
297	34	46.6	219	2	S32120	RNA-directed DNA p	370	34	46.6	337	2	S26863	glyceralddehyde-3-p
298	34	46.6	219	2	S32086	RNA-directed DNA p	371	34	46.6	338	2	T47218	glyceralddehyde-3-p
299	34	46.6	219	2	S32133	RNA-directed DNA p	372	34	46.6	339	2	G64041	glyceralddehyde-3-p
300	34	46.6	219	2	S32081	RNA-directed DNA p	373	34	46.6	342	1	S64042	glyceralddehyde-3-p
301	34	46.6	219	2	S32139	RNA-directed DNA p	374	34	46.6	342	1	HLHUC4	MHC class I histoc
302	34	46.6	219	2	S32078	RNA-directed DNA p	375	34	46.6	342	1	AC3299	pseudouridylate sy
303	34	46.6	219	2	S32066	RNA-directed DNA p	376	34	46.6	350	2	D71273	probable glycerald
304	34	46.6	219	2	S32126	RNA-directed DNA p	377	34	46.6	350	2	C87536	transcription regu
305	34	46.6	219	2	S32074	RNA-directed DNA p	378	34	46.6	350	2	AG0879	A/G-specific adeni
306	34	46.6	219	2	S32053	RNA-directed DNA p	379	34	46.6	350	2	A40647	methylational-indepe
307	34	46.6	219	2	S32088	RNA-directed DNA p	380	34	46.6	354	2	C83577	hypothetical prote
308	34	46.6	219	2	S32072	RNA-directed DNA p	381	34	46.6	355	2	I80169	class I histocompa
309	34	46.6	219	2	S32070	RNA-directed DNA p	382	34	46.6	361	2	A48445	glyceralddehyde-3-p
310	34	46.6	219	2	S32087	RNA-directed DNA p	383	34	46.6	361	2	T29571	hypothetical prote
311	34	46.6	219	2	S32076	RNA-directed DNA p	384	34	46.6	363	2	JH0542	class I histocompa
312	34	46.6	219	2	S32092	RNA-directed DNA p	385	34	46.6	363	2	S07113	probable trehalose
313	34	46.6	219	2	S32160	RNA-directed DNA p	386	34	46.6	365	2	D95881	MHC class I histoc
314	34	46.6	219	2	S32054	RNA-directed DNA p	387	34	46.6	365	2	S42823	MHC class I histoc
315	34	46.6	219	2	S32135	RNA-directed DNA p	388	34	46.6	366	2	B37028	ubiquinol-cytochro
316	34	46.6	219	2	S32159	RNA-directed DNA p	389	34	46.6	366	2	D81381	ubiquinol-cytochro
317	34	46.6	219	2	S32159	RNA-directed DNA p	390	34	46.6	379	2	I48133	ubiquinol-cytochro
318	34	46.6	219	2	S32069	RNA-directed DNA p	391	34	46.6	379	2	I48132	ubiquinol-cytochro
319	34	46.6	219	2	S32063	RNA-directed DNA p	392	34	46.6	379	2	I48180	ubiquinol-cytochro
320	34	46.6	219	2	S32128	RNA-directed DNA p	393	34	46.6	379	2	I48180	ubiquinol-cytochro
321	34	46.6	219	2	S32085	RNA-directed DNA p	394	34	46.6	379	2	I48134	ubiquinol-cytochro

395 34 46.6 385 2 B82921 serine/threonine k
396 34 46.6 399 1 B70936 probable serine/th
397 34 46.6 401 2 G87552 conserved hypochet
398 34 46.6 407 2 T36404 probable monooxyge
399 34 46.6 410 2 H86290 hypochetical prote
400 34 46.6 412 2 G90362 hypochetical prote
401 34 46.6 413 2 JC6512 methyltransferase
402 34 46.6 417 2 E97789 amgG protein [impo
403 34 46.6 424 2 B83442 probable MFS trans
404 34 46.6 431 2 T50177 probable peptide m
405 34 46.6 432 2 AF2801 hypochetical prote
406 34 46.6 432 2 H97580 hypochetical prote
407 34 46.6 442 2 B84232 hypochetical prote
408 34 46.6 446 2 AC2812 conserved hypochet
409 34 46.6 446 2 D97590 hypochetical prote
410 34 46.6 463 2 G90657 hypochetical prote
411 34 46.6 463 2 G85508 hypochetical prote
412 34 46.6 465 2 S41644 polyadenylate-bind
413 34 46.6 470 2 AD0888 SufI protein [impo
414 34 46.6 470 2 B91116 suppressor of ftsI
415 34 46.6 470 2 B85961 suppressor of ftsI
416 34 46.6 472 2 G65088 sufi protein precu
417 34 46.6 472 2 T75505 hypochetical prote
418 34 46.6 472 2 A13357 multidrug resistan
419 34 46.6 499 2 AC2058 cell death suppres
420 34 46.6 500 2 JC7658 dipeptidyl-peptida
421 34 46.6 518 2 T48143 flavonoid 3',5'-hy
422 34 46.6 519 2 C86160 hypochetical prote
423 34 46.6 526 2 F83166 hypochetical prote
424 34 46.6 541 2 AP2483 hypochetical prote
425 34 46.6 559 2 B47175 reverse transcript
426 34 46.6 559 2 B47175 reverse transcript
427 34 46.6 568 2 E90364 hypochetical prote
428 34 46.6 571 2 AI0506 probable sulfatase
429 34 46.6 584 2 C75364 probable long-chai
430 34 46.6 594 2 G83878 L-lactate permease
431 34 46.6 610 2 B82436 penicillin-binding
432 34 46.6 623 2 T48859 disease resistance
433 34 46.6 629 2 T16767 hypochetical prote
434 34 46.6 629 2 T16767 hypochetical prote
435 34 46.6 629 2 T16767 hypochetical prote
436 34 46.6 652 2 T16582 hypochetical prote
437 34 46.6 687 2 T27421 hypochetical prote
438 34 46.6 708 2 B83884 hypochetical prote
439 34 46.6 802 2 T45642 FtsH metalloprotei
440 34 46.6 868 2 C81200 acetonate hydratase
441 34 46.6 868 2 H81775 acetonate hydratase
442 34 46.6 882 1 WMBE05 gene UL5 protein -
443 34 46.6 885 1 JC5085 replication licens
444 34 46.6 902 2 T01688 pol polyprotein -
445 34 46.6 908 2 B69435 signal-transducing
446 34 46.6 912 2 S33980 pol polyprotein -
447 34 46.6 924 2 E71476 alanine-tRNA ligase
448 34 46.6 955 2 T48515 hypochetical prote
449 34 46.6 1002 1 GNLJND HIV-1 retropepsin
450 34 46.6 1002 2 S54378 pol polyprotein -
451 34 46.6 1003 1 GNLVIV HIV-1 retropepsin
452 34 46.6 1003 1 B44001 HIV-1 retropepsin
453 34 46.6 1003 1 GNLVW2 HIV-1 retropepsin
454 34 46.6 1003 2 T09440 pol polyprotein -
455 34 46.6 1008 2 S38003 translation elonga
456 34 46.6 1012 1 GNLVIV HIV-1 retropepsin
457 34 46.6 1015 1 GNLVH3 HIV-1 retropepsin
458 34 46.6 1027 1 GNLJ81 HIV-1 retropepsin
459 34 46.6 1054 1 GNLJG5 HIV-1 retropepsin
460 34 46.6 1061 1 GNLJG4 HIV-1 retropepsin
461 34 46.6 1123 2 T28139 PKA protein kinase
462 34 46.6 1168 2 T30935 reverse transcript
463 34 46.6 1197 2 D82696 hypochetical prote
464 34 46.6 1205 2 B95845 conserved hypochet
465 34 46.6 1220 1 DJBEC3 DNA-directed DNA p
466 34 46.6 1220 2 T42573 DNA-directed DNA p
467 34 46.6 1271 2 T43269 microcystin synthe

468 34 46.6 1411 2 S55123 hypochetical prote
469 34 46.6 1487 2 B82083 glutamate synthase
470 34 46.6 1620 2 B83261 conserved hypochet
471 34 46.6 1707 2 A33526 collagen alpha 2(I
472 34 46.6 1774 2 S13178 6-methylsalicylic
473 34 46.6 2359 2 B96832 hypochetical prote
474 34 46.6 3036 2 T18995 hypochetical prote
475 34 46.6 3433 1 GNLVUV genome polyprotein
476 33.5 45.9 3433 1 GNLVUV genome polyprotein
477 33.5 45.9 209 2 B87627 hypochetical prote
478 33.5 45.9 265 2 B87511 conserved hypochet
479 33.5 45.9 344 2 T05437 hypochetical prote
480 33.5 45.9 348 2 H85256 hypochetical prote
481 33.5 45.9 355 2 S62692 L-amino acid oxida
482 33.5 45.9 412 2 AF0868 probable benzozate
483 33.5 45.9 471 2 C82825 UDP-N-acetylmuram
484 33.5 45.9 480 2 T36822 probable transcrip
485 33.5 45.9 519 2 D82536 conserved hypochet
486 33.5 45.9 584 2 D84264 hypochetical prote
487 33.5 45.9 685 2 JC6331 rho-type guanine e
488 33.5 45.9 686 2 B87490 NADH dehydrogenase
489 33 45.2 39 2 T37555 homeobox - human (
490 33 45.2 62 2 C89895 50S ribosomal prot
491 33 45.2 75 2 S03065 ser protein - huma
492 33 45.2 85 2 AH1010 conserved hypochet
493 33 45.2 94 2 T48710 hypochetical prote
494 33 45.2 114 2 E71171 hypochetical prote
495 33 45.2 142 2 AF0961 heat shock protein
496 33 45.2 150 2 G95348 nitric-oxide reduc
497 33 45.2 154 2 AC0496 heat shock protein
498 33 45.2 159 2 S32327 coat protein - Ses
499 33 45.2 169 2 S26011 hypochetical prote
500 33 45.2 170 2 B83871 hypochetical prote
501 33 45.2 190 2 A82343 conserved hypochet
502 33 45.2 193 2 T35847 probable carbonic
503 33 45.2 195 2 AI2549 hypochetical prote
504 33 45.2 196 2 AI0502 molybdopterin bios
505 33 45.2 206 2 AD1898 urease accessory p
506 33 45.2 207 2 C36361 hypochetical prote
507 33 45.2 208 2 AD2808 conserved hypochet
508 33 45.2 209 1 XURTMC methylated-DNA-lpr
509 33 45.2 209 2 A83849 alpha-ribazole-5'-
510 33 45.2 213 2 AC3389 uracil-DNA glycosy
511 33 45.2 213 2 AH2253 hypochetical prote
512 33 45.2 214 1 ASLJG5 vif protein - simi
513 33 45.2 214 1 ASLJSM vif protein - simi
514 33 45.2 217 2 B97587 hypochetical prote
515 33 45.2 218 2 JC7220 nuclear protein SR
516 33 45.2 219 2 S32095 RNA-directed DNA p
517 33 45.2 219 2 A46097 GPI-anchor biosynt
518 33 45.2 224 2 F70814 probable malonyl c
519 33 45.2 229 2 JC7219 nuclear protein SR
520 33 45.2 253 2 JQ2255 triose-phosphate i
521 33 45.2 257 2 T51696 3-methyl-2-oxobuta
522 33 45.2 260 2 B81194 biotin synthesis p
523 33 45.2 260 2 B83833 hypochetical prote
524 33 45.2 262 2 I40221 diVb protein - Ba
525 33 45.2 264 2 AG2095 hypochetical prote
526 33 45.2 264 2 AH2202 hypochetical prote
527 33 45.2 269 2 S73999 hypochetical prote
528 33 45.2 274 2 H72521 probable thiazole
529 33 45.2 288 2 S39888 SMR2 protein - Pod
530 33 45.2 293 2 AF0617 probable DNA methy
531 33 45.2 295 2 E75366 glutamyl-tRNA synt
532 33 45.2 300 2 S69185 glyceraldehyde-3-p
533 33 45.2 302 2 S75227 hypochetical prote
534 33 45.2 306 2 AE2885 transcription regu
535 33 45.2 306 2 B97661 probable transcrip
536 33 45.2 308 2 B87059 probable pseudouri
537 33 45.2 308 2 A70761 hypochetical prote
538 33 45.2 308 2 S72886 hypochetical prote
539 33 45.2 313 2 T35826 probable dehydroge
540 33 45.2 315 2 B84937 cysteine synthase

541	33	45.2	317	2	G83544	probable transcrip	614	33	45.2	513	1	S50216	translation initia
542	33	45.2	321	2	T42591	gene 48 protein -	615	33	45.2	529	2	E87259	amido-phosphoribosy
543	33	45.2	330	2	T64905	probable sugar tra	616	33	45.2	533	2	S43526	hypothetical prote
544	33	45.2	330	2	G85723	probable transport	617	33	45.2	535	2	S76564	ribitol kinase [im
545	33	45.2	330	2	B90894	probable transpor	618	33	45.2	536	2	AH3087	D-ribulose 1,4-bet
546	33	45.2	331	1	DEUTC	glyceraldhyde-3-p	619	33	45.2	536	2	A99199	cellulose 1,4-bet
547	33	45.2	332	1	DEJNGI	glyceraldhyde-3-p	620	33	45.2	540	2	S41942	hypothetical prote
548	33	45.2	332	1	T39602	glyceraldhyde-3-p	621	33	45.2	540	2	S76869	phytoene dehydrog
549	33	45.2	334	2	G71504	glyceraldhyde-3-p	622	33	45.2	541	2	F84187	transcription fact
550	33	45.2	335	2	JC7529	glyceraldhyde-3-p	623	33	45.2	542	1	A54963	glutaryl-tRNA synt
551	33	45.2	335	2	B72053	glyceraldhyde 3-p	624	33	45.2	551	2	D69282	glucose-6-phosphat
552	33	45.2	335	2	B86568	glyceraldhyde-3-p	625	33	45.2	560	2	S41908	hypothetical prote
553	33	45.2	335	2	H81662	glyceraldhyde 3-p	626	33	45.2	567	2	T16105	hypothetical prote
554	33	45.2	335	2	S59579	glyceraldhyde-3-p	627	33	45.2	568	2	T05218	hypothetical prote
555	33	45.2	335	2	T40252	glyceraldhyde-3-p	628	33	45.2	582	2	T16104	hypothetical prote
556	33	45.2	335	2	H70524	probable fabH prot	629	33	45.2	594	2	B82994	probable cell cycl
557	33	45.2	336	1	D69231	probable X-Pro dip	630	33	45.2	600	2	F75424	hypothetical prote
558	33	45.2	336	2	T40235	glyceraldhyde-3-p	631	33	45.2	610	2	S41315	hypothetical integ
559	33	45.2	336	2	H70693	hypothetical prote	632	33	45.2	618	2	A10171	NADH2 dehydrogenas
560	33	45.2	336	2	H73389	hypothetical prote	633	33	45.2	625	2	T40742	calcium binding PW
561	33	45.2	337	2	S26974	glyceraldhyde-3-p	634	33	45.2	633	2	B70946	probable protein k
562	33	45.2	337	2	S26973	glyceraldhyde-3-p	635	33	45.2	634	2	JC4248	phage transposase
563	33	45.2	337	2	T27635	homeobox protein c	636	33	45.2	648	1	H69878	hypothetical prote
564	33	45.2	338	2	AH3321	transcription regu	637	33	45.2	696	2	A91247	polysphosphate kina
565	33	45.2	343	2	A10581	DNA polymerase III	638	33	45.2	698	2	T39050	transcription-rela
566	33	45.2	343	2	B85564	DNA polymerase III	639	33	45.2	739	2	T45429	topoisomerase IV c
567	33	45.2	343	2	F90713	DNA polymerase III	640	33	45.2	742	2	E70673	myosin-Ic - mouse
568	33	45.2	343	2	A45251	DNA-directed DNA p	641	33	45.2	775	2	A83402	ATP-dependent heli
569	33	45.2	347	1	S01340	D-amino-acid oxida	642	33	45.2	807	2	H75634	probable ATP-depen
570	33	45.2	347	1	S01340	hypothetical prote	643	33	45.2	822	2	E75523	probable large ATP
571	33	45.2	355	2	AG1877	hypothetical prote	644	33	45.2	838	2	B83150	probable large ATP
572	33	45.2	364	2	C84187	hypothetical prote	645	33	45.2	861	2	T36381	probable large ATP
573	33	45.2	366	2	T42972	hypothetical prote	646	33	45.2	862	2	T36380	alanyl-tRNA synthet
574	33	45.2	383	2	A71390	ubiquinol-cytochro	647	33	45.2	875	2	H81739	hypothetical prote
575	33	45.2	383	2	T20572	hypothetical prote	648	33	45.2	885	2	G91080	hypothetical prote
576	33	45.2	384	2	S51796	vasodilator-estimul	649	33	45.2	899	2	H85925	hypothetical prote
577	33	45.2	386	2	A83025	probable acyl-CoA	650	33	45.2	913	2	T35718	hypothetical prote
578	33	45.2	389	2	B89096	corrinoid/iron-sul	651	33	45.2	1000	2	T30280	C50C3.2 protein -
579	33	45.2	389	2	B93910	probable serine/th	652	33	45.2	1009	2	S44621	probable large ATP
580	33	45.2	393	2	JC5614	RNAs protein - rat	653	33	45.2	1010	2	T36383	DNA-directed DNA p
581	33	45.2	399	2	S75030	hypothetical prote	654	33	45.2	1012	1	DDBE68	probable DNA-direc
582	33	45.2	401	2	AF2418	hypothetical prote	655	33	45.2	1012	2	T44185	DNA polymerase [im
583	33	45.2	408	2	A86652	LPS biosynthesis p	656	33	45.2	1012	2	T43998	myosin I beta - hu
584	33	45.2	408	2	E70127	flagellar motor sw	657	33	45.2	1028	2	A59253	myosin heavy chain
585	33	45.2	408	2	T50876	hypothetical membr	658	33	45.2	1028	2	S41749	myosin I heavy cha
586	33	45.2	409	2	T51126	hypothetical prote	659	33	45.2	1028	2	S37146	recombination acti
587	33	45.2	411	2	T22290	hypothetical prote	660	33	45.2	1073	2	I51055	isoleucyl-tRNA syn
588	33	45.2	414	2	S75052	hypothetical prote	661	33	45.2	1078	2	F75407	probable arabinosy
589	33	45.2	416	2	E69374	N-ethylmellamine ch	662	33	45.2	1094	2	F70697	exodeoxyribonuclea
590	33	45.2	416	2	E87286	3-deoxy-D-manno-oc	663	33	45.2	1121	2	G64103	HIV-1 retropepsin
591	33	45.2	420	2	E75378	probable valine-py	664	33	45.2	1124	1	GNLJFP	HIV-1 retropepsin
592	33	45.2	420	2	C83246	probable binding p	665	33	45.2	1124	2	B45557	pol polyprotein -
593	33	45.2	430	2	AG0531	cell cycle protein	666	33	45.2	1124	2	S23820	hypothetical prote
594	33	45.2	431	2	A81150	histidyl-tRNA synt	667	33	45.2	1127	2	T21635	ribonucleoside-dip
595	33	45.2	432	2	A83060	hypothetical prote	668	33	45.2	1127	2	T21635	DNA-directed DNA p
596	33	45.2	435	2	B74218	conserved hypotet	669	33	45.2	1137	1	WMBE1	hypothetical prote
597	33	45.2	440	2	I49681	glyceraldhyde-3-p	670	33	45.2	1143	2	A69465	hypothetical prote
598	33	45.2	448	2	B45438	myosin I beta, NM	671	33	45.2	1165	2	T21636	hypothetical prote
599	33	45.2	451	2	S77599	probable coproporp	672	33	45.2	1256	2	S14556	asparagine-rich pr
600	33	45.2	451	2	AF3053	conserved hypotet	673	33	45.2	1319	1	C43735	bcsC protein - Ace
601	33	45.2	451	2	G98232	hypothetical prote	674	33	45.2	1339	2	T47841	hypothetical prote
602	33	45.2	458	2	B75413	major facilitator	675	33	45.2	1404	1	A48196	DNA polymerase III
603	33	45.2	467	2	A57627	p55 erythrocyte me	676	33	45.2	1451	2	D64203	protein-tyrosine k
604	33	45.2	469	2	B70607	probable cysS prot	677	33	45.2	1616	2	T17884	S-layer protein -
605	33	45.2	473	2	C86949	probable cysteinyl	678	33	45.2	1733	2	D70887	polyketide synthas
606	33	45.2	480	2	T24087	hypothetical prote	679	33	45.2	1784	2	E86921	hypothetical prote
607	33	45.2	485	2	E83620	probable amidase p	680	33	45.2	1882	2	S73484	toxin-like outer m
608	33	45.2	487	2	A12692	conserved hypotet	681	33	45.2	2399	2	H71879	hypothetical prote
609	33	45.2	487	2	E97474	BH3939 hypotetica	682	33	45.2	3430	1	GNWVW	genome polyprotein
610	33	45.2	499	2	H83254	probable MFS trans	683	33	45.2	7576	2	T17428	FK506 polyketide s
611	33	45.2	505	2	S68518	tub protein, brain	684	33	45.2	202	2	S56267	probable membrane
612	33	45.2	510	2	S41943	cellulose 1,4-beta	685	32.5	44.5	202	2	T03526	blub protein - Rho
613	33	45.2	511	2	S44716	cellulose 1,4-beta	686	32.5	44.5	219	2	F87444	dethiobiotin synth

687	32.5	44.5	245	2	A84279	hypothetical prote	760	32	43.8	269	2	E83240	hypothetical prote
688	32.5	44.5	417	2	AC0546	probable metabolic	761	32	43.8	272	2	T35231	hypothetical prote
689	32.5	44.5	678	2	D82415	exoribonuclease II	762	32	43.8	278	2	S06416	DNA-directed DNA p
690	32.5	44.5	1736	2	T00391	hypothetical prote	763	32	43.8	281	2	F84127	hypothetical prote
691	32	43.8	62	2	S01004	trypsin inhibitor	764	32	43.8	284	2	AH0146	hypothetical prote
692	32	43.8	62	2	A95051	ribosomal protein	765	32	43.8	281	2	H86159	probable sulfura
693	32	43.8	62	2	F97921	50S ribosomal prot	766	32	43.8	286	2	H87292	enoyl-CoA hydratase
694	32	43.8	84	2	D83528	hypothetical prote	767	32	43.8	288	2	S73016	H ⁺ -transporting tw
695	32	43.8	84	2	A23301	hypothetical prote	768	32	43.8	288	2	T10566	6-phosphogluconate
696	32	43.8	85	2	B91147	hypothetical prote	769	32	43.8	293	2	T35157	6-phosphogluconate
697	32	43.8	85	2	F85992	hypothetical prote	770	32	43.8	294	2	I40701	glyceralddehyde-3-p
698	32	43.8	85	2	C65120	hypothetical prote	771	32	43.8	294	2	I41220	glyceralddehyde-3-p
699	32	43.8	93	2	I47624	hypothetical 10.0	772	32	43.8	294	2	I41118	glyceralddehyde-3-p
700	32	43.8	93	2	E70967	Ig heavy chain V-V	773	32	43.8	294	2	I41119	glyceralddehyde-3-p
701	32	43.8	99	2	AH3429	hypothetical prote	774	32	43.8	294	2	I41222	glyceralddehyde-3-p
702	32	43.8	104	2	E82797	transposase BMER14	775	32	43.8	294	2	I41221	glyceralddehyde-3-p
703	32	43.8	105	2	D72546	conserved hypothet	776	32	43.8	294	2	I41227	glyceralddehyde-3-p
704	32	43.8	111	2	H95908	hypothetical prote	777	32	43.8	294	2	A96155	hypothetical prote
705	32	43.8	112	2	S75152	conserved hypothet	778	32	43.8	299	2	H87049	hypothetical prote
706	32	43.8	114	2	S22559	hypothetical prote	779	32	43.8	299	2	B83888	transcription regu
707	32	43.8	117	2	AB2757	Ig heavy chain V r	780	32	43.8	300	2	S75892	hypothetical prote
708	32	43.8	123	2	T49336	hypothetical prote	781	32	43.8	302	2	A70865	hypothetical prote
709	32	43.8	125	1	A46315	hypothetical prote	782	32	43.8	302	2	F95387	probable transfera
710	32	43.8	125	1	F81197	E4 protein - human	783	32	43.8	305	2	AH0863	protein imported
711	32	43.8	132	2	S15618	holo-(acyl-carrier	784	32	43.8	305	2	F85932	regulatory protein
712	32	43.8	133	1	UC2003	E4 protein - human	785	32	43.8	305	2	I41065	positive regulator
713	32	43.8	142	2	B82223	NADH ubiquinone ox	786	32	43.8	305	2	D91087	glycine cleavage s
714	32	43.8	144	2	G83228	hypothetical prote	787	32	43.8	308	2	A82391	positive regulator
715	32	43.8	145	2	AG2706	probable transcrip	788	32	43.8	309	2	F83434	transcription regu
716	32	43.8	147	2	T38472	conserved hypothet	789	32	43.8	310	2	JC7853	translocation prot
717	32	43.8	147	2	A70928	hypothetical prote	790	32	43.8	312	2	S76507	L-fucose-specific
718	32	43.8	150	2	AF3634	hypothetical prote	791	32	43.8	315	2	H83276	hypothetical prote
719	32	43.8	153	2	F70950	nitric-oxide reduc	792	32	43.8	315	2	F91450	probable lipase PA
720	32	43.8	158	2	D72305	hypothetical prote	793	32	43.8	323	2	G83461	probable major hea
721	32	43.8	167	2	S52220	hypothetical prote	794	32	43.8	323	2	E95864	hypothetical prote
722	32	43.8	170	1	NM02	cobinamide kinase	795	32	43.8	328	2	H90889	hypothetical prote
723	32	43.8	170	2	AB2294	2S albumin 2 precu	796	32	43.8	328	2	G64901	ABC-type transport
724	32	43.8	181	2	G70692	hypothetical prote	797	32	43.8	328	2	A85728	hypothetical prote
725	32	43.8	183	2	AC0767	hypothetical prote	798	32	43.8	329	1	DRVKGL	hypothetical prote
726	32	43.8	183	2	S15302	hDTP-4-dehydrotham	799	32	43.8	330	2	T36944	hypothetical prote
727	32	43.8	186	2	G97498	hypothetical prote	800	32	43.8	331	1	DEECG3	glyceralddehyde-3-p
728	32	43.8	195	1	SAVLDS	delta large antige	801	32	43.8	331	2	S57279	glyceralddehyde-3-p
729	32	43.8	199	2	S72728	probable L-a-amino	802	32	43.8	331	2	S57281	glyceralddehyde-3-p
730	32	43.8	200	1	XUSMEG	hDTP-4-dehydrotham	803	32	43.8	331	2	AG0711	glyceralddehyde-3-p
731	32	43.8	206	1	NBHUIB	platelet glycoprot	804	32	43.8	331	2	S57280	glyceralddehyde-3-p
732	32	43.8	208	2	JN0638	platelet glycoprot	805	32	43.8	331	2	D85788	glyceralddehyde-3-p
733	32	43.8	210	2	G83082	hypothetical prote	806	32	43.8	331	2	H90939	glyceralddehyde-3-p
734	32	43.8	210	2	G85729	hypothetical prote	807	32	43.8	332	1	DEBYG1	glyceralddehyde-3-p
735	32	43.8	214	2	S97989	vif protein - simi	808	32	43.8	332	1	DEBYG2	glyceralddehyde-3-p
736	32	43.8	214	2	T11561	vif protein - simi	809	32	43.8	332	1	DEBYG3	glyceralddehyde-3-p
737	32	43.8	215	2	D70692	hypothetical prote	810	32	43.8	333	1	DEKZGR	glyceralddehyde-3-p
738	32	43.8	218	2	H83475	probable transcrip	811	32	43.8	333	2	JC6310	glyceralddehyde-3-p
739	32	43.8	219	2	T18541	mofB protein precu	812	32	43.8	334	2	D82803	glyceralddehyde-3-p
740	32	43.8	220	2	D82139	conserved hypothet	813	32	43.8	334	2	A10262	glyceralddehyde-3-p
741	32	43.8	221	2	D87489	transcription regu	814	32	43.8	334	2	G86723	exodeoxyribonuclea
742	32	43.8	221	2	B36881	MM2-reactive phos	815	32	43.8	335	2	S43339	glyceralddehyde-3-p
743	32	43.8	222	2	A81896	hypothetical prote	816	32	43.8	335	2	T12046	glyceralddehyde-3-p
744	32	43.8	222	2	F81124	conserved hypothet	817	32	43.8	337	2	JC5023	UMP-staic acid tr
745	32	43.8	222	2	B90888	hypothetical prote	818	32	43.8	338	2	S26976	glyceralddehyde-3-p
746	32	43.8	225	2	G83371	probable amino aci	819	32	43.8	338	2	JN0452	glyceralddehyde-3-p
747	32	43.8	228	2	A12779	nitroreductase [im	820	32	43.8	338	2	A44132	D-aspartate oxidas
748	32	43.8	231	2	D82555	heme ABC transport	821	32	43.8	338	2	C83570	hypothetical prote
749	32	43.8	235	2	T19328	hypothetical prote	822	32	43.8	340	2	T09663	hypothetical prote
750	32	43.8	236	2	H75275	DNA-binding respon	823	32	43.8	340	2	AD0701	tetrathionate redu
751	32	43.8	238	2	D97559	probable nitroredu	824	32	43.8	341	2	T35426	probable oxidoredu
752	32	43.8	244	2	A82316	hypothetical prote	825	32	43.8	347	2	B64069	yceG protein homol
753	32	43.8	249	2	A70840	hypothetical prote	826	32	43.8	350	2	F75259	glutanyl-tRNA redu
754	32	43.8	252	2	C84522	22 kDa peroxisomal	827	32	43.8	353	2	B72738	hypothetical prote
755	32	43.8	252	2	TJ1439	probable cobyr a	828	32	43.8	355	2	C83354	hypothetical prote
756	32	43.8	254	2	T35994	probable ABC-type	829	32	43.8	359	2	D83103	hypothetical prote
757	32	43.8	263	2	T35686	phosphatidylglycer	830	32	43.8	363	2	T45920	hypothetical prote
758	32	43.8	267	2	AC3133	hypothetical prote	831	32	43.8	365	2	B83829	glycine oxidase (s
759	32	43.8	268	2	B42424	chitinase (EC 3.2.	832	32	43.8	368	2	G65119	hypothetical 40.4

833	32	43.8	371	2	A88520	41.8K hypothetical	906	32	43.8	575	2	A11996	hypothetical prote
834	32	43.8	373	2	E70338	probable aspartate	907	32	43.8	583	2	A13098	hypothetical prote
835	32	43.8	376	2	C87596	glycosyl hydrolase	908	32	43.8	591	2	G90988	serine/threonine p
836	32	43.8	377	2	A87288	glutamate 5-kinase	909	32	43.8	592	2	T42078	alkylglycerone-pho
837	32	43.8	378	2	D64385	aspartate transami	910	32	43.8	597	2	JC5829	probable membrane
838	32	43.8	380	2	S70964	pkn5 protein - Myx	911	32	43.8	601	2	T36323	flagellar hook-ass
839	32	43.8	386	2	H87480	conserved hypotet	912	32	43.8	606	2	G64659	prSD protein (U891
840	32	43.8	387	2	T01210	glucose-6-phosphat	913	32	43.8	606	2	H98187	hypothetical prote
841	32	43.8	388	2	G70729	hypothetical prote	914	32	43.8	610	2	T16761	probable exported
842	32	43.8	393	2	B85932	probable transport	915	32	43.8	615	2	A10615	hypothetical prote
843	32	43.8	394	2	C82439	peptide methionine	916	32	43.8	615	2	H90754	probable amidase [
844	32	43.8	396	1	XNECD	aspartate transami	917	32	43.8	615	2	F85618	probable amidase y
845	32	43.8	396	2	AD0616	aspartate aminotra	918	32	43.8	615	2	D64832	yCB protein precu
846	32	43.8	396	2	A85619	aspartate aminotra	919	32	43.8	615	2	A82838	ycbB topoisomerase
847	32	43.8	396	2	C90755	conserved hypotet	920	32	43.8	622	2	JC7973	synleurin - human
848	32	43.8	396	2	T35254	carbamoyl-phosphat	921	32	43.8	634	2	T00054	hypothetical prote
849	32	43.8	399	2	B82723	probable transport	922	32	43.8	634	1	W1L558	E1 protein - human
850	32	43.8	399	2	F91146	probable transpor	923	32	43.8	646	2	A85834	probable chaperoni
851	32	43.8	403	2	AD0748	tyrosine-specific	924	32	43.8	646	2	E64973	hypothetical prote
852	32	43.8	409	2	T47118	platelet glycoprot	925	32	43.8	648	2	T43337	polo-like kinase-1
853	32	43.8	411	1	I55604	hypothetical prote	926	32	43.8	651	2	T25953	hypothetical prote
854	32	43.8	414	2	C75461	probable membrane	927	32	43.8	659	2	E86313	cytochrome c-type
855	32	43.8	415	2	S53075	probable trna meth	928	32	43.8	660	1	S54746	phosphatidylinosit
856	32	43.8	415	2	T38324	transcription init	929	32	43.8	664	1	PC4002	2',3'-cyclic-nucle
857	32	43.8	416	1	JC4952	aspartate transami	930	32	43.8	681	1	H82059	hypothetical prote
858	32	43.8	418	2	S46315	interferon consens	931	32	43.8	682	2	T41680	hypothetical prote
859	32	43.8	424	2	A35861	regulator protein	932	32	43.8	684	2	F70810	hypothetical prote
860	32	43.8	424	2	S44506	interferon consens	933	32	43.8	686	2	T08919	probable bacteriop
861	32	43.8	425	2	A45064	probable C4-dicarb	934	32	43.8	691	2	T44543	hypothetical prote
862	32	43.8	427	2	B83001	hypothetical prote	935	32	43.8	709	2	F75584	DOA1 protein - Yea
863	32	43.8	428	2	B71403	hypothetical prote	936	32	43.8	715	2	S38051	penicillin-binding
864	32	43.8	429	2	D96027	probable nitritotr	937	32	43.8	719	2	A83800	hypothetical prote
865	32	43.8	452	2	A11877	proline/betaine tr	938	32	43.8	722	2	G86746	conserved hypotet
866	32	43.8	454	2	C82682	glutamate-cysteine	939	32	43.8	725	2	A11544	hypothetical prote
867	32	43.8	456	2	T06136	aspartate transami	940	32	43.8	727	2	T08920	hypothetical prote
868	32	43.8	461	2	H64636	proline/betaine tr	941	32	43.8	738	2	C95936	hypothetical prote
869	32	43.8	469	2	A99656	beta-fructofuranos	942	32	43.8	738	2	D75598	hypothetical prote
870	32	43.8	473	2	T06167	hypothetical prote	943	32	43.8	761	2	T09052	hypothetical prote
871	32	43.8	475	2	T01352	hypothetical prote	944	32	43.8	763	2	E96571	transport protein
872	32	43.8	482	2	D75346	glutamyl-L-RRNA(Gln)	945	32	43.8	765	2	T09574	carbon-monoxide de
873	32	43.8	489	2	A11276	multidrug-efflux t	946	32	43.8	775	2	B64319	hypothetical prote
874	32	43.8	489	2	A11639	multidrug-efflux t	947	32	43.8	783	2	T30644	primosomal replica
875	32	43.8	493	2	AB0451	conserved hypotet	948	32	43.8	798	2	C98069	1-phosphatidylinos
876	32	43.8	493	2	AD1398	drug-export protei	949	32	43.8	801	1	T52538	dimethylsulfoxide
877	32	43.8	493	2	AG1773	drug-export protei	950	32	43.8	808	2	E64914	probable oxidoredu
878	32	43.8	498	2	AH3456	glutamate synthase	951	32	43.8	808	2	E90915	probable oxidoredu
879	32	43.8	499	2	B85507	hypothetical prote	952	32	43.8	830	1	A57060	serine/threonine-s
880	32	43.8	502	2	D86578	S/T protein kinase	953	32	43.8	837	2	A97557	dimethylglycine de
881	32	43.8	502	2	G72045	s/t protein kinase	954	32	43.8	837	2	AD2777	outer membrane pro
882	32	43.8	503	2	E72078	probable sodium-tr	955	32	43.8	837	2	F75525	cell surface antig
883	32	43.8	503	2	A86544	NADH (ubiquinone)	956	32	43.8	846	2	P75525	nikB protein - Esc
884	32	43.8	503	2	B81690	probable sodium-tr	957	32	43.8	875	1	A57080	DNA-directed DNA p
885	32	43.8	503	2	C71535	probable NADH (ubi	958	32	43.8	899	2	B38529	hypothetical prote
886	32	43.8	504	2	A9467	occludin - chicken	959	32	43.8	911	2	T37659	hypothetical prote
887	32	43.8	504	2	T32175	hypothetical prote	960	32	43.8	914	2	S18942	exonuclease ABC c
888	32	43.8	515	2	D90048	hypothetical prote	961	32	43.8	916	2	H73272	SNR2/rad54 helicase
889	32	43.8	525	2	C69794	glutamate synthase	962	32	43.8	916	2	G75417	poly(A) polymerase
890	32	43.8	531	1	XPFS7A	site-specific DNA-	963	32	43.8	942	2	S75598	probable transcrip
891	32	43.8	532	1	E69821	multidrug resistan	964	32	43.8	942	2	T34847	probable cadmium-t
892	32	43.8	535	2	T070580	probable UDP-N-ace	965	32	43.8	943	2	T08987	hypothetical prote
893	32	43.8	537	2	AE2454	two-component sens	966	32	43.8	951	2	S32954	male-specific leth
894	32	43.8	538	2	H96008	hypothetical prote	967	32	43.8	952	2	T32954	hypothetical prote
895	32	43.8	540	2	B45665	adult-specific 61.	968	32	43.8	955	2	S52959	phage lambda-relat
896	32	43.8	540	2	B87350	hypothetical prote	969	32	43.8	971	2	T24866	probable membrane
897	32	43.8	544	2	T38469	conserved hypotet	970	32	43.8	971	2	T14968	hypothetical prote
898	32	43.8	552	1	E57987	cytochrome c-type	971	32	43.8	972	2	C94781	hypothetical prote
899	32	43.8	552	2	D86101	hypothetical prote	972	32	43.8	984	2	A83324	probable sensor/re
900	32	43.8	552	2	H91260	cytochrome c-type	973	32	43.8	984	2	A83324	ksr protein - frui
901	32	43.8	559	2	A10135	DNA repair protein	974	32	43.8	1003	2	T13856	oxoglutarate dehyd
902	32	43.8	560	2	S65148	tRNA adenyllyltrans	975	32	43.8	1014	1	DEBY	klcloho protein - r
903	32	43.8	564	2	E70394	ABC transporter (h	976	32	43.8	1014	1	JEO333	formate dehydrogen
904	32	43.8	567	2	S69778	adhesin Ape5-1 pre	977	32	43.8	1016	1	S40838	
905	32	43.8	573	2	I60247	SEC23 protein homo	978	32	43.8				

979 32 43.8 1016 2 D91231
980 32 43.8 1019 2 T00117
981 32 43.8 1094 2 C70612
982 32 43.8 1102 2 S55100
983 32 43.8 1170 2 S03308
984 32 43.8 1175 2 T46124
985 32 43.8 1190 2 T38636
986 32 43.8 1194 2 DJBE28
987 32 43.8 1203 2 H87687
988 32 43.8 1204 2 T18812
989 32 43.8 1235 1 DJBE16
990 32 43.8 1235 1 DJBEAN
991 32 43.8 1235 1 DJBEH7
992 32 43.8 1235 1 DJBEK1
993 32 43.8 1235 1 DJBEV1
994 32 43.8 1240 1 DJBEZ1
995 32 43.8 1261 2 G83162
996 32 43.8 1268 2 S52781
997 32 43.8 1299 2 AH2090
998 32 43.8 1333 2 S30356
999 32 43.8 1369 2 T03104
1000 32 43.8 1462 1 S32437

ALIGNMENTS

RESULT 1

A36925
transcription activator LysR-type Cbbr - Xanthobacter flavus
C/Species: Xanthobacter flavus
C/Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 09-Jul-2004
C/Accession: A36925; S13578; S35408
R/van den Bergh, E.R.E.; Dijkhuizen, L.; Meijer, W.G.
J. Bacteriol. 175, 6097-6104, 1993
A/Title: Cbbr, a LysR-type transcriptional activator, is required for expression of the
A/Reference number: A36925; MUID:94012468; PMID:8407781
A/Accession: A36925
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-333 <VAN>
A/Cross-references: UNIPROT:P25545; UNIPARC:UPI0000127169; EMBL:Z22705; NID:G297851; PID
R/Meijer, W.G.; Arnberg, A.C.; Enequist, H.G.; Terpetra, P.; Lidstrom, M.E.; Dijkhuizen,
Mol. Gen. Genet. 225, 320-330, 1991
A/Title: Identification and organization of carbon dioxide fixation genes in Xanthobacter
A/Reference number: S13573; MUID:91172133; PMID:1900916
A/Accession: S13578
A/Molecule type: DNA
A/Residues: 1-150 <MEI>
A/Cross-references: UNIPARC:UPI00001788AC; EMBL:X17252
C/Genetics:

A/Gene: cbbR

A/Start codon: GTG

C/Superfamily: transcription activator LysR-type

C/Keywords: DNA binding; transcription regulation

Query Match 67.8%; Score 49.5; DB 2; Length 333;
Best Local Similarity 66.7%; Pred. No. 0.99;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 IEQ-PTLRQWLAARA 14

Db 264 VEGLPVVRQWLAARA 278

RESULT 2

AG0147
Probable membrane protein YPO1203 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C/Accession: AG0147
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AG0147
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-296 <KUR>
A/Cross-references: UNIPROT:Q8ZGS7; UNIPARC:UPI00000DC87B; GB:AL590842; PIDN:CAC90042.1;
C/Genetics:
A/Gene: YPO1203
C/Superfamily: hypothetical protein ydeD

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PTLRQWLAARA 14

Db 66 PTLRQWLAARA 76

RESULT 3

D70601
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galu [similarity] - Mycobacteri
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C/Accession: D70601
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Hollroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: D70601
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-306 <COL>
A/Cross-references: UNIPROT:O05576; UNIPARC:UPI00000CCAD8; GB:Z94752; GB:AL123456; NID:G
A/Experimental source: strain H37Rv
C/Genetics:

A/Gene: galU

C/Superfamily: UDP-glucose pyrophosphorylase

C/Keywords: nucleotidyltransferase

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 3.5;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAR 13

Db 290 GPDLLRWLVAR 300

RESULT 4

T23485
hypothetical protein K08F4.11 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T23485
R/Hembry, C.
submitted to the EMBL Data Library, January 1996
A/Reference number: Z19746
A/Accession: T23485
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-200 <WIL>
A/Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN00022;
A/Experimental source: clone K08F4
C/Genetics:
A/Gene: CESP:K08F4.11
A/Map position: 4

A;Introns: 45/1; 76/1; 111/3
C;Superfamily: glutathione transferase

Query Match 60.3%; Score 44; DB 2; Length 200;
Best Local Similarity 61.5%; Pred. No. 4.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAAR 13
||| ||| ||| ||| |||
Db 183 IETPKLKEWLAKR 195

RESULT 5
T37464
probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37464
R;Tave, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsens, K.
submitted to the EMBL Data Library, June 1997
A;Description: Paraquat mediates differential gene expression in C. elegans.
A;Reference number: Z21702
A;Accession: T37464
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-207 <TA>
A;Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AAB65419.
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: GST3
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 60.3%; Score 44; DB 2; Length 207;
Best Local Similarity 61.5%; Pred. No. 5.1;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAAR 13
||| ||| ||| ||| |||
Db 190 IETPKLKEWLAKR 202

RESULT 6
SYECBB
2,3-dihydroxybenzoate-[carrier protein] ligase (EC 6.2.1.1-) entE - Escherichia coli (str
N;Alternate names: 2,3-dihydroxybenzoate-AMP ligase [misnomer]; dihydroxybenzoic acid-ac
C;Species: Escherichia coli
C;Date: 31-Dec-1989 #sequence_revision 21-Nov-1997 #text_change 09-Jul-2004
C;Accession: H64792; A48308; A32047; I41058; S08076
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64792
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-536 <BLAT>
A;Cross-references: UNIPROT:P10378; UNIPARC:UPI0000129FD2; GB:AE000165; GB:U00096; NID:9
A;Experimental source: strain K-12, substrain MG1655
R;Stad, J.F.; Elkins, M.F.; Earhart, C.F.
FEMS Microbiol. Lett. 59, 15-19, 1989
A;Title: Nucleotide sequence of the Escherichia coli entE gene.
A;Reference number: A48308; MUID:89290355; PMID:2525505
A;Note: in MedLine 89290355 this citation is erroneously given as volume 50 rather than
A;Accession: A48308
A;Molecule type: DNA
A;Residues: 1-368, 'ECRRKSTAAAR', 379-536 <STA>
A;Cross-references: UNIPARC:UPI000016F103; GB:M27490; EMBL:X15058; NID:g41345; PIDN:CAA3
R;Liu, J.; Duncan, K.; Waleh, C.T.
J. Bacteriol. 171, 791-798, 1989
A;Title: Nucleotide sequence of a cluster of Escherichia coli enterobactin biosynthesis
A;Reference number: A91904; MUID:89123155; PMID:2521622

A;Accession: A32047
A;Molecule type: DNA
A;Residues: 393-536 <LIU>
A;Cross-references: UNIPARC:UPI000016F4EA; GB:M24148; NID:g304949; PIDN:AAA16101.1; PID:
C;Comment: The enzymatic steps in the condensation of L-serine and 2,3-dihydroxybenzoic
ty is based on its recognized homology with 4-coumarate-CoA ligase and by analogy with t
C;Comment: The formation of 2,3-dihydroxybenzoyl-AMP has been observed. The rapid reacti
carrier protein) to release AMP, has also been observed.
C;Genetics:
A;Gene: entE
A;Map position: 14 min
C;Function:
A;Description: catalyzes the formation of 2,3-dihydroxybenzoyl-[carrier protein], AMP an
A;Pathway: enterobactin biosynthesis
A;Note: this is one component of a membrane-bound multienzyme complex that catalyzes the
for transport into the cell
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
C;Keywords: acid-thiol ligase; enterobactin biosynthesis; membrane-associated complex
F;69-526/Domain: acetate-CoA ligase homology <ACL>

Query Match 60.3%; Score 44; DB 1; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
:: |||||::||
Db 521 VDKKQLRQWLASRA 534

RESULT 7
E85558
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
C;Accession: E85558
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85558
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-536 <STO>
A;Cross-references: UNIPARC:UPI00001656E8; GB:AE005174; NID:g12513487; PIDN:AAG54929.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: entE
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
:: |||||::||
Db 521 VDKKQLRQWLASRA 534

RESULT 8
A99708
2,3-dihydroxybenzoate-AMP ligase [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A99708
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99708
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-536 <HAY>
A;Cross-references: UNIPROT:Q8XEV3; UNIPARC:UPI00000129FD1; GB:BA000007; PIDN:BAB34056.1;
A;Experimental source: strain O157:H7, substrain RMD 050952
C;Genetics:
C;Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 60.3%; Score 44; DB 2; Length 536;
Best Local Similarity 57.1%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
:: |||||::
Db 521 VDKKQLRQWLASRA 534

RESULT 9
G71337
probable dimethyladenosine transferase (ksa) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 31-Dec-2004
C;Accession: G71337
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: G71337
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-285 <COL>
A;Cross-references: UNIPROT:O83357; UNIPARC:UPI0000012E0CA; GB:AE001213; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0337
C;Superfamily: dimethyladenosine transferase (rRNA adenosine dimethyltransferase)

Query Match 58.9%; Score 43; DB 2; Length 285;
Best Local Similarity 64.3%; Pred. No. 10;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 14
||| |||||
Db 98 IEGDVLOQWRAAAA 111

RESULT 10
B71325
conserved hypothetical protein TP0421 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: B71325
R;Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: B71325
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-683 <COL>
A;Cross-references: UNIPROT:O83436; UNIPARC:UPI000000C0A71; GB:AE001220; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0421

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAARA 13

||| |||||
Db 89 IEGAAALHONGAAR 101

RESULT 11
E84853
hypothetical protein At2g242400 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: E84853
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <STO>
A;Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI0000017A02D; GB:AE002093; NID:g4567312; PI
C;Genetics:
A;Gene: At2g242400
A;Map position: 2

Query Match 57.5%; Score 42; DB 2; Length 473;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEGPTLRQWL 10
:|||::|
Db 343 VEGETIREWL 352

RESULT 12
T11560
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)
C;Species: simian immunodeficiency virus SIVsm
A;Variety: strain E543
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11560
R;Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.; N
J. Virol. 71, 1608-1620, 1997
A;Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficer
A;Reference number: Z17285; MUID:97131152; PMID:8995688
A;Accession: T11560
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1019 <HIR>
A;Cross-references: UNIPROT:P89154; UNIPARC:UPI0000105470; EMBL:U72748; NID:gi695908; PI
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; immunodeficiency

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGPTLRQW 9
||| |||||
Db 184 EGPKLRQW 191

RESULT 13
F91171
probable phosphopantetheinyltransferase [imported] - Escherichia coli (strain O157:H7, s
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F91171
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91171
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <HAY>
A;Cross-references: UNIPROT:Q8X5U4; UNIPARC:UPI000012534F; GB:BA000007; PIDN:BAB37765.1;
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EC84342

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAARA 14
:||:|||||
Db 27 QGPRRERWLAGRA 39
:|||||

RESULT 14
F86017
probable phosphopantetheinyltransferase [imported] - Escherichia coli (strain O157:H7, a
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F86017
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F86017
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <STO>
A;Cross-references: UNIPROT:Q8X5U4; UNIPARC:UPI000012534F; GB:AE005174; NID:gi2518155; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4867

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAARA 14
:||:|||||
Db 27 QGPRRERWLAGRA 39
:|||||

RESULT 15
S47694
hypothetical 21.8K protein (ftsY-nika intergenic region) - Escherichia coli (strain K-12
N;Alternate names: hypothetical protein o195
C;Species: Escherichia coli
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47694; F65144
R;Plunkett, G.
submitted to the EMBL Data Library, March 1994
A;Reference number: S47666
A;Accession: S47694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-195 <PLU>
A;Cross-references: UNIPROT:P37623; UNIPARC:UPI0000125350; EMBL:U000039; NID:g466582; PID
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65144
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-195 <BLAT>

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAARA 14
:||:|||||
Db 27 QGPRRERWLAGRA 39
:|||||

RESULT 16
E87575
ABC transporter, ATP-binding protein CC2634 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87575
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: UNIPROT:Q9A535; UNIPARC:UPI000000C77B8; GB:AE005673; NID:gi13424211; P
C;Genetics:
A;Gene: CC2634

Query Match 56.2%; Score 41; DB 2; Length 249;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAAR 13
:|||||
Db 76 QAPTLPWLSAR 87
:|||||

RESULT 17
T45453
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galu [similarity] - Mycobacteri
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 05-Oct-2004
C;Accession: T45453
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z22967
A;Accession: T45453
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-306 <JAM>
A;Cross-references: UNIPROT:Q9Z5G1; UNIPARC:UPI00000D438E; EMBL:AL035500; PIDN:CAB36696.
A;Experimental source: cosmid L373
C;Genetics:
A;Note: galu
A;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidylyltransferase

Query Match 56.2%; Score 41; DB 2; Length 306;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAR 13
:|||||
Db 290 GPDLRNVLVER 300
:|||||

RESULT 18

A;Cross-references: UNIPARC:UPI0000125350; GB:AE000423; GB:U00096; NID:gi1789880; PIDN:RA
A;Experimental source: strain K-12, substrain MGL655
C;Genetics:
A;Gene: yhhU

Query Match 56.2%; Score 41; DB 2; Length 195;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAARA 14
:||:|||||
Db 27 QGPRRERWLAGRA 39
:|||||

C24430
glyceraldehyde-3-phosphate dehydrogenase (NADP) (phosphorylating) (EC 1.2.1.13) C, cytosol
C/Species: Nicotiana glauca (Common tobacco)
C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C/Accession: C24430
R/Shih, M.C.; Lazar, G.; Goodman, H.M.
Cell 47, 73-80, 1986
A/Title: Evidence in favor of the symbiotic origin of chloroplasts: primary structure and
A/Reference number: A90888; MUID:87002494; PMID:3757034
A/Accession: C24430
A/Molecule type: mRNA
A/Residues: 1-326 <SHI>
A/Cross-references: UNIPROT:P09094; UNIPARC:UPI000012AE7D; GB:M14419; NID:g170240; PIDN:
C/Genetics:
A/Gene: gapC
C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: cytosol; NADP; oxidative phosphorylation; oxidoreductase
Query Match 56.2%; Score 41; DB 2; Length 326;
Best Local Similarity 35.7%; Pred. No. 26; Mismatches 5; Indels 0; Gaps 0;
Matches 5; Conservative 5
Qy 1 IEQPTLRQWLAARA 14
Db 179 VDGPSMKDMRGRA 192
RESULT 19
DEPZG
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - parsley
C/Species: Petroselinum crispum (parsley)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: S18484
R/Martin, W.; Gierl, A.; Saedler, H.
Nature 339, 46-48, 1989
A/Title: Molecular evidence for pre-Cretaceous angiosperm origins.
A/Reference number: S17991
A/Accession: S18484
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: mRNA
A/Residues: 1-336 <MAR>
A/Cross-references: UNIPROT:P26519; UNIPARC:UPI000012AE76; EMBL:X60344; NID:g20548; PIDN:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1991
C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F/4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F/153,180/Active site: Cys, His #status predicted
Query Match 56.2%; Score 41; DB 1; Length 336;
Best Local Similarity 35.7%; Pred. No. 27; Mismatches 5; Indels 0; Gaps 0;
Matches 5; Conservative 5
Qy 1 IEQPTLRQWLAARA 14
Db 189 VDGPSMKDMRGRA 202
RESULT 20
A35080
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - common ice plant
C/Species: Mesembryanthemum crystallinum (common ice plant)
C/Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 09-Jul-2004
C/Accession: A35080
R/Ostrem, J.A.; Vernon, D.M.; Bohnert, H.J.
J. Biol. Chem. 265, 3497-3502, 1990
A/Title: Increased expression of a gene coding for NAD:glyceraldehyde-3-phosphate dehydrogenase
allium.
A/Reference number: A35080; MUID:90154012; PMID:2303458
A/Accession: A35080
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-337 <OST>
A/Cross-references: UNIPROT:P17878; UNIPARC:UPI000012AE73; GB:J05223; NID:g167263; PIDN:

C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: oxidoreductase
Query Match 56.2%; Score 41; DB 2; Length 337;
Best Local Similarity 35.7%; Pred. No. 27; Mismatches 5; Conservative 5; Indels 0; Gaps 0;
Matches 5; Conservative 5
Qy 1 IEQPTLRQWLAARA 14
Db 190 VDGPSMKDMRGRA 203
RESULT 21
DEIS3C
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - wh
N/Alternate names: triosephosphate dehydrogenase
C/Species: Sinapis alba (white mustard)
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C/Accession: A24796
R/Martin, W.; Cerff, R.
Eur. J. Biochem. 159, 323-331, 1986
A/Title: Prokaryotic features of a nucleus-encoded enzyme. cDNA sequences for chloroplast
A/Reference number: A24796; MUID:87004643; PMID:3530755
A/Accession: A24796
A/Molecule type: mRNA
A/Residues: 1-338 <MAR>
A/Cross-references: UNIPROT:P04796; UNIPARC:UPI000016DF6D; GB:X04301; NID:g21142; PIDN:CF
C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F/2-338/Product: glyceraldehyde-3-phosphate dehydrogenase #status experimental <MAT>
F/7-37/Region: beta-alpha-beta NAD nucleotide-binding fold
F/156,183/Active site: Cys, His #status predicted
Query Match 56.2%; Score 41; DB 1; Length 338;
Best Local Similarity 35.7%; Pred. No. 27; Mismatches 5; Conservative 5; Indels 0; Gaps 0;
Matches 5; Conservative 5
Qy 1 IEQPTLRQWLAARA 14
Db 192 VDGPSMKDMRGRA 205
RESULT 22
JQ1287
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12), cytosolic - Arab
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: JQ1287; JS0614
R/Shih, M.C.; Heinrich, P.; Goodman, H.M.
Gene 104, 133-138, 1991
A/Title: Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cyto
A/Reference number: JQ1285; MUID:92009205; PMID:1916285
A/Accession: JQ1287
A/Molecule type: DNA
A/Residues: 1-338 <SHI>
A/Cross-references: UNIPROT:P25858; UNIPARC:UPI000016DB30; GB:M64119; NID:g166709; PIDN:J
A/Molecule type: mRNA
A/Residues: 1-338 <SHI1>
A/Cross-references: UNIPARC:UPI000016DB30; GB:M64116; NID:g166705; PIDN:AAA32794.1; PIDN:
A/Experimental source: leaf
C/Genetics:
A/Gene: gapC
A/Map position: 3 0.0cM
C/Introns: 2/1; 12/1; 45/3; 84/2; 117/3; 167/2; 187/1; 267/2
C/Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C/Keywords: cytosol; oxidoreductase
Query Match 56.2%; Score 41; DB 2; Length 338;
Best Local Similarity 35.7%; Pred. No. 27; Mismatches 5; Conservative 5; Indels 0; Gaps 0;
Matches 5; Conservative 5
Qy 1 IEQPTLRQWLAARA 14

```

Db      192 VDGPSMKDWRGGR 205
      ::|:::|
RESULT 23
B95325 conserved hypothetical protein SMA0937 [imported] - Sinorhizobium meliloti (strain 1021)
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95325
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: B95325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:CROSS-references: UNIPROT:Q92ZH9; UNIPARC:UPI00000CB0F3; GB:AE006469; PIDN:AAK65164.1;
F:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0937
A:Genome: plasmid

      Query Match      56.2%; Score 41; DB 2; Length 719;
      Best Local Similarity 46.2%; Pred. No. 58;
      Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEQPTLROWLAAR 13
      ::|:::|
Db      71 LDDPEVRQWLTA 83

RESULT 24
A97501 topoisomerase iv chain a [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C:Accession: A97501
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A97501
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-750 <KUR>
A:CROSS-references: UNIPROT:Q8UG82; UNIPARC:UPI00000D1A6D; GB:AE007869; PIDN:AAK86962.1;
C:Genetics:
A:Gene: AGR_C_2144
A:Map position: circular chromosome
C:Superfamily: Type II topoisomerase, subunit A; phage T4 DNA topoisomerase (ATP-hydroly

      Query Match      56.2%; Score 41; DB 2; Length 750;
      Best Local Similarity 66.7%; Pred. No. 61;
      Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      3 GPTLROWLAARA 14
      |:::|
Db      721 GEELREWLADRA 732

RESULT 25
AE2719 topoisomerase IV subunit A parC [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C:Accession: AE2719
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, P.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer, Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2719
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-750 <KUR>
A:CROSS-references: UNIPROT:Q8UG82; UNIPARC:UPI00000D1A6D; GB:AE008688; PIDN:AAAL42171.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: parC
A:Map position: circular chromosome
C:Superfamily: Type II topoisomerase, subunit A; phage T4 DNA topoisomerase (ATP-hydroly

      Query Match      56.2%; Score 41; DB 2; Length 750;
      Best Local Similarity 66.7%; Pred. No. 61;
      Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      3 GPTLROWLAARA 14
      |:::|
Db      721 GEELREWLADRA 732

RESULT 26
S25204 srMx protein - Streptomyces ambofaciens
C:Species: Streptomyces ambofaciens
C>Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C:Accession: S25204; S21599
R:Geistlich, M.; Losick, R.; Turner, J.R.; Rao, R.N.
Mol. Microbiol. 6, 2019-2029, 1992
A:Title: Characterization of a novel regulatory gene governing the expression of a polyk
A:Reference number: S25202; MUID:92374852; PMID:1508047
A:Accession: S25204
A:Molecule type: DNA
A:Residues: 1-239 <GEI>
A:CROSS-references: UNIPROT:Q00510; UNIPARC:UPI00000BEAE2; EMBL:X63451; NID:g46699; PIDN:
C:Genetics:
A:Gene: srMx
F:39-139/Domain: bioC homology <BIOC>

      Query Match      54.8%; Score 40; DB 2; Length 239;
      Best Local Similarity 50.0%; Pred. No. 28;
      Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      1 IEQPTLROWLAARA 14
      :|:::|
Db      63 VSGLELSEWMAARA 76

RESULT 27
S27491 hypothetical protein A - Bacillus firmus
C:Species: Bacillus firmus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S27491
R:Quirk, P.G.; Krulwich, T.A.
submitted to the EMBL Data Library, October 1991
A:Reference number: S27490
A:Accession: S27491
A:Status: preliminary
A:Molecule type: DNA
```

```
A;Residues: 1-463 <QUI>
A;Cross-references: UNIPROT:P30267; UNIPARC:UPI00001396AB; GB:L02548; EMBL:M74194; NID:9

Query Match      54.8%; Score 40; DB 2; Length 463;
Best Local Similarity 61.5%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  2 EGPTLRQWLAAARA 14
Db  296 EGKTSRQWALERA 308

RESULT 28
A81958
probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: A81958
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jegels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-530 <PAR>
A;Cross-references: UNIPROT:Q9WE3; UNIPARC:UPI00000C497F; GB:AL162753; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0414

Query Match      54.8%; Score 40; DB 2; Length 530;
Best Local Similarity 72.7%; Pred. No. 62;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  1 IEPTLRQWLA 11
Db  190 IEMPVLRPWLA 200

RESULT 29
E81015
ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: E81015
R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: E81015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <TET>
A;Cross-references: UNIPROT:Q9UXI9; UNIPARC:UPI00000C4849; GB:AE002552; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2026

Query Match      54.8%; Score 40; DB 2; Length 531;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  1 IEPTLRQWLA 11
Db  191 IEMPVLRPWLA 201

RESULT 30
```

```
S30484
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30484
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30484
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
C;Superfamily: pol polyprotein

Query Match      54.8%; Score 40; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy  1 IEPTLRQW 9
Db  29 MDGPKLRQW 37

RESULT 31
S30483
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C;Accession: S30483
R;Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A;Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A;Reference number: S30460
A;Accession: S30483
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-656 <GAO>
A;Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111
C;Superfamily: pol polyprotein

Query Match      54.8%; Score 40; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy  1 IEPTLRQW 9
Db  29 MDGPKLRQW 37

RESULT 32
A39707
erythrocyte membrane band 4.2 protein - human
N;Alternate names: pallidin
N;Contains: erythrocyte membrane band 4.2 protein, long splice form; erythrocyte membrane
C;Species: Homo sapiens (man)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: A39707; A34865; A34883
R;Korngren, C.; Cohen, C.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 4840-4844, 1991
A;Title: Organization of the gene for human erythrocyte membrane protein 4.2: structural
A;Reference number: A39707; MUID:91271288; PMID:2052563
A;Accession: A39707
A;Molecule type: DNA
A;Residues: 1-721 <KOR1>
A;Cross-references: UNIPROT:P16452; UNIPARC:UPI00001608BD; GB:L06519; NID:9306738; PIDN:
R;Sung, L.A.; Chien, S.; Chang, L.S.; Lambert, K.; Blas, S.A.; Bouhassira, E.E.; Nagel,
Proc. Natl. Acad. Sci. U.S.A. 87, 955-959, 1990
A;Title: Molecular cloning of human protein 4.2: a major component of the erythrocyte me
A;Reference number: A34865; MUID:90138995; PMID:1689063
A;Accession: A34865
```

A;Molecule type: mRNA
A;Residues: 1-364, 'KRGLPC', 371-379, 'H', 381-405, 'L', 407-721 <SUN1>
A;Cross-references: UNIPARC:UPI000016AE3E; GB:M30647; NID:g189433; PIDN:AAA36401.1; PID:
A;Accession: B34865
A;Molecule type: mRNA
A;Residues: 1-3, 34-364, 'KRGLPC', 371-379, 'H', 381-405, 'L', 407-721 <SUN2>
A;Cross-references: UNIPARC:UPI000016AE3F; GB:M30646; NID:g189435; PIDN:AAA36402.1; PID:
A;Experimental source: isolate Sickle cell patient; cell type reticulocyte
A;Note: parts of this sequence were determined by protein sequencing
R;Korsgren, C.; Lawler, J.; Lambert, S.; Speicher, D.; Cohen, C.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 613-617, 1990
A;Title: Complete amino acid sequence and homologies of human erythrocyte membrane prote
A;Reference number: A34883; MUID:90138879; PMID:2300550
A;Accession: A34883
A;Molecule type: mRNA
A;Residues: 1-3, 34-721 <ROR2>
A;Cross-references: UNIPARC:UPI000016A88B; GB:M29399; NID:g182083; PIDN:AAA35798.1; PID:
A;Comment: This protein is a major constituent of the erythrocyte membrane. It apparentl
C;Genetics:
A;Gene: GDB:EPB42; PA
A;Cross-references: GDB:127385; OMIM:177070
A;Map position: 15q15-15q15
C;Superfamily: protein-glutamine gamma-glutamyltransferase
C;Keywords: alternative splicing; blocked amino end; glycoprotein; lipoprotein; myristyl
P:2-721/Product: erythrocyte membrane band 4.2 protein, long splice form #status predic
P:2-3, 34-721/Product: erythrocyte membrane band 4.2 protein, short splice form #status p
F:298-316/Domain: transmembrane #status predicted <TRM>
F:518-520/Region: cell attachment (R-G-D) motif
F:2, 420, 447, 529, 604, 705/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:103, 420, 447, 529, 604, 705/Binding site: carbohydrate (Ser) (covalent) (by CAMP-dependent kinase) #status predic
F:278/Binding site: phosphate (Ser) (covalent)
Query Match 54.8%; Score 40; DB 2; Length 721;
Best Local Similarity 70.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 PTLQWLAR 13
DB 280 PILQWLTR 289
RESULT 33
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N;Alternate names: protein F5E19_190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:QSPU7; UNIPARC:UPI000008A945F; EMBL:AL391147
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2
A;Note: F5E19_190
Query Match 54.8%; Score 40; DB 2; Length 1123;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 IEQPTLRQWLAA 12
DB 200 VQPTKRWLSS 211
RESULT 34
CGH02B

collagen alpha 2(IV) chain precursor - human
N;Alternate names: procollagen alpha 2(IV) chain
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A32024; S00007; S02624; S00246; S17678; S16911; B32117; S16877; S00165; S39
R;Hostikka, S.L.; Tryggvason, K.
J. Biol. Chem. 263, 19488-19493, 1988
A;Title: The complete primary structure of the alpha2 chain of human type IV collagen an
A;Reference number: A32024; MUID:89066769; PMID:3198637
A;Accession: A32024
A;Molecule type: mRNA
A;Residues: 1-1712 <HOS1>
A;Cross-references: UNIPROT:P08572; UNIPARC:UPI0000126D42; EMBL:J04210; EMBL:X05610; GB:M20753; NID:g2995
R;Hostikka, S.L.; Kurkinen, M.; Tryggvason, K.
FEBS Lett. 216, 281-286, 1987
A;Title: Nucleotide sequence coding for the human type IV collagen alpha-2 chain cDNA re
ated region.
A;Reference number: S00007; MUID:87219158; PMID:3582677
A;Accession: S00007
A;Molecule type: mRNA
A;Residues: 1254-1398, 'V', 1400-1712 <HOS2>
A;Cross-references: UNIPARC:UPI0000173BE6; EMBL:J04210; EMBL:X05610; GB:M20753; NID:g2995
A;Note: 1399-1le was also found
R;Hostikka, S.L.; Tryggvason, K.
FEBS Lett. 224, 297-305, 1987
A;Title: Extensive structural differences between genes for the alpha(1) and alpha(2) ch
A;Reference number: S02624; MUID:88083553; PMID:2826228
A;Accession: S02624
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1347-1350; 1377-1383; 1426-1432; 1465-1471; 1529-1535; 1625-1630 <HOS3>
A;Cross-references: UNIPARC:UPI0000142E66; UNIPARC:UPI0000173BE7; UNIPARC:UPI0000173BE8;
A;Note: Complete nucleotide sequence not shown
R;Brazel, D.; Pollner, R.; Oberbauer, I.; Kuehn, K.
Eur. J. Biochem. 172, 35-42, 1988
A;Title: Human basement membrane collagen (type IV): the amino acid sequence of the alph
A;Reference number: S00246; MUID:88151998; PMID:3345760
A;Accession: S00246
A;Molecule type: mRNA
A;Residues: 1-682, 'G', 684-1043 <BRA>
A;Cross-references: UNIPARC:UPI0000173BEC; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PII
R;Oberbauer, I.
submitted to the EMBL Data Library, June 1987
A;Reference number: S17678
A;Accession: S17678
A;Molecule type: mRNA
A;Residues: 1-470, 'P', 472-682, 'G', 684-1043 <OBE>
A;Cross-references: UNIPARC:UPI000016A709; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PII
R;Poeschl, E.; Pollner, R.; Kuehn, K.
EMBO J. 7, 2687-2695, 1988
A;Title: The genes for the alpha(IV) and alpha2(IV) chains of human basement membrane
A;Reference number: S02738; MUID:89030632; PMID:2846280
A;Accession: S16911
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <POE>
A;Cross-references: UNIPARC:UPI000016AGF3; EMBL:X12784; GB:M36963; NID:g30072; PIDN:CAA
R;Soininen, R.; Huotari, M.; Hostikka, S.L.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 263, 17217-17220, 1988
A;Title: The structural genes for alpha1 and alpha2 chains of human type IV collagen ar
A;Reference number: A92690; MUID:89034231; PMID:3182844
A;Accession: B32117
A;Molecule type: DNA
A;Residues: 1-33 <SO11>
A;Cross-references: UNIPARC:UPI000016AGF3; EMBL:J04217; EMBL:J05039; NID:g180759; PIDN:
R;Soininen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Tryggvason, K.
J. Biol. Chem. 264, 13565-13571, 1989
A;Title: Structural organization of the gene for the alpha-1 chain of human type IV coll
A;Reference number: S16876; MUID:89340433; PMID:2701944
A;Accession: S16877
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-33 <SO12>

A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; NID:G180759; PIDN:AAA53097.1; PI
A;Note: this sequence was submitted to the EMBL Data Library, October 1988
R;Siebold, B.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 168, 569-575, 1987
A;Title: Construction of a model for the aggregation and cross-linking region (7S domain
is region.
A;Reference number: S00165; MUID:88029476; PMID:3117548
A;Accession: S00165
A;Molecule type: protein
A;Residues: 37-247 <SIE1>
A;Cross-references: UNIPARC:UPI0000173BED
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-Gly
R;Eble, J.A.; Golbik, R.; Mann, K.; Kuehn, K.
EMBO J. 12, 4795-4802, 1993
A;Title: The alpha-1-beta-1 integrin recognition site of the basement membrane collagen
A;Reference number: S39614; MUID:94038963; PMID:8223488
A;Accession: S39615
A;Molecule type: protein
A;Residues: 407-570 <EEL>
A;Cross-references: UNIPARC:UPI0000173BEE
R;MacWright, R.S.; Benson, V.A.; Lovello, K.T.; van der Rest, M.; Fietzek, P.P.
Biochemistry 22, 4940-4948, 1983
A;Title: Isolation and characterization of pepsin-solubilized human basement membrane (t
A;Reference number: S16910; MUID:84053346; PMID:6416291
A;Accession: S16912
A;Molecule type: protein
A;Residues: 490-492, 'X', 494-496; 675-677, 'G', 679-680, 'G', 682, 684-685, 'P' <MAC>
A;Cross-references: UNIPARC:UPI0000173BEF; UNIPARC:UPI0000173BF0
A;Experimental source: placenta
R;Glanville, R.W.; Rauter, A.
Hoppe-Seyler's Z. Physiol. Chem. 362, 943-951, 1981
A;Title: Pepsin fragments of human placental basement-membrane collagens showing interrup
A;Reference number: S16908; MUID:82005835; PMID:6792033
A;Accession: S58517
A;Molecule type: protein
A;Residues: 490-492, 'X', 494-501, 'P', 503-507; 952-957, 'X', 959-966, 'X', 968; 984-986, 'X', 988-
81-1185 <GLA>
A;Cross-references: UNIPARC:UPI0000173BF1; UNIPARC:UPI0000173BF2; UNIPARC:UPI0000173BF3;
R;Killen, P.D.; Francomano, C.A.; Yamada, Y.; Modi, W.S.; O'Brien, S.J.
Hum. Genet. 77, 318-324, 1987
A;Title: Partial structure of the human alpha-2(IV) collagen chain and chromosomal local
A;Reference number: S01450; MUID:88085168; PMID:3692475
A;Accession: S01450
A;Molecule type: mRNA
A;Residues: 1040, 'L', 1042-1398, 'V', 1400-1418, 'M', 1420-1635, 'V', 1637-1712 <KIL>
A;Cross-references: UNIPARC:UPI0000072E68; EMBL:M24766; NID:G537328; PIDN:AAA52043.1; PI
R;Siebold, B.; Deutzmann, R.; Kuehn, K.
Eur. J. Biochem. 176, 617-624, 1988
A;Title: The arrangement of intra- and intermolecular disulfide bonds in the carboxyterm
A;Reference number: S02550; MUID:89005112; PMID:2844531
A;Accession: S02550
A;Molecule type: protein
A;Residues: 1480-1535; 1545-1614; 1617-1662, 'H', 1664-1700, 'G', 1705-1708; 1710-1712 <SIE2>
A;Cross-references: UNIPARC:UPI0000173BF6; UNIPARC:UPI0000173BF7; UNIPARC:UPI0000173BF8;
A;Note: the sequence from Fig. 7 is inconsistent with that shown in Fig. 11 in having 17
R;Myers, J.C.; Howard, P.S.; Jelen, A.M.; Dion, A.S.; Macarak, E.J.
J. Biol. Chem. 262, 9231-9238, 1987
A;Title: Duplication of type IV collagen COOH-terminal repeats and species-specific exp
A;Reference number: A27114; MUID:87250571; PMID:2439508
A;Accession: B27114
A;Molecule type: mRNA
A;Residues: 1486-1574, 'I', 1576-1712 <MYE>
A;Cross-references: UNIPARC:UPI0000173BF8; EMBL:J02760; NID:G180425; PIDN:AAA58422.1; PI
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL4A2
A;Map position: 13q34-13q34
A;Introns: 15/2; 33/3; 1347/1; 1380/1; 1468/1; 1532/1; 1527/3 #status incomplete
A;Note: the alpha 1(IV) and alpha 2(IV) chain genes are encoded on opposite strands with
C;Complex: Type IV collagen is a heterotrimer of two alpha 1(IV) chains (see PIR:CGHU4B)
domains (with disulfide and desmosine cross-links), dimeric associations among trimer ca

rupted helical domain (with disulfide and desmosine cross-links).
C;Function:
A;Description: structural component of basement membrane
C;Superfamily: collagen alpha 1(IV) chain
K;Keywords: basement membrane; cell binding; coiled coil; extracellular matrix; glycoprotein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-57/Domain: amino-terminal nonhelical, NH1 <NH1>
F;58-1485/Region: interrupted helical
F;58-1485/Region: cell attachment (R-G-D) motif
F;582-364/Region: cell attachment (R-G-D) motif
F;784-786/Region: cell attachment (R-G-D) motif
F;868-870/Region: cell attachment (R-G-D) motif
F;889-891/Region: cell attachment (R-G-D) motif
F;970-972/Region: cell attachment (R-G-D) motif
F;1069-1071/Region: cell attachment (R-G-D) motif
F;1128-1230/Region: cell attachment (R-G-D) motif
F;1452-1454/Region: cell attachment (R-G-D) motif
F;1486-1712/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1495-1593/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1603-1708/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;42,47,51,53,137,483,485/Disulfide bonds: interchain #status predicted
F;57,87,90,102,165,168,225,239,242/Binding site: carboxylate (Lys) (covalent) #status predicted
F;57/Modified site: 5-hydroxylysine (Lys) #status atypical
F;63,75,96,114,120,123,132,150,159,186,189,198,201,213,216,219,496,499,955,964,1103,1115,
F;87,90,102,165,168,225,239,242/Modified site: 5-hydroxylysine (Lys) #status experimental
F;138/Binding site: carboxylate (Asn) (covalent) #status experimental
F;209/Modified site: 4-hydroxyproline (Pro) #status atypical
F;661-681/Disulfide bonds: #status predicted
F;1275/Binding site: carboxylate (Asn) (covalent) #status predicted
F;1504-1590,1537-1593/Disulfide bonds: (or 1504-1593, 1537-1590) #status experimental
F;1549-1555,1658-1665/Disulfide bonds: #status experimental
F;1612-1705,1646-1708/Disulfide bonds: (or 1612-1708, 1646-1705) #status experimental
Query Match 54.8%; Score 40; DB 1; Length 1712;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 IEQPTLRQWL 10
Db 8 VAGPALRRRL 17
RESULT 35
A84326
hypothetical protein Vngl740c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84326
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; La
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84326
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <STO>
A;Cross-references: UNIPROT:Q9HP97; UNIPARC:UPI000006396C; GB:AB004437; NID:G10581200; P
C;Genetics:
A;Gene: VNG1740C
Query Match 54.1%; Score 39.5; DB 2; Length 325;
Best Local Similarity 64.3%; Pred. No. 46;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Qy 1 IEQPTLRQWLAAARA 14
Db 14 LEGPA-AAWLAARA 26
RESULT 36

S74539
hypothetical protein slr0740 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74539
O: Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
R.; Kaneo, T.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KAN>
A:Cross-references: UNIPROT:P72684; UNIPARC:UPI00000C09C4; EMBL:D90899; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: *Synechocystis* hypothetical protein slr0740

Query Match 53.4%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTLQWLA 11
| | | | |
Db 29 POLQWLA 36
| | | | |

RESULT 37
S46354
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)
C:Species: simian immunodeficiency virus SIVagm
A:Variety: isolate SABD37
C:Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C:Accession: S46354
R: Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;
EMBO J. 13, 2935-2947, 1994
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A:Reference number: S46335; MUID:94298785; PMID:8026477
A:Accession: S46354
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-217 <JIN>
A:Cross-references: UNIPARC:UPI000010A53B; EMBL:U04018; NID:G466250; PIDN:AAA21512.1; PI
A:Experimental source: isolate SABD37; *sabaeus* monkey
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
: | | | | |
Db 86 DGPRLRQW 93
: | | | | |

RESULT 38
I40327
baf protein - *Bordetella pertussis*
C:Species: *Bordetella pertussis*
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I40327; S70669
R: DeShazer, D.; Wood, G.E.; Friedman, R.L.
J. Bacteriol. 177, 3801-3807, 1995
A:Title: Identification of a *Bordetella pertussis* regulatory factor required for transcri
A:Reference number: I40327; MUID:95325323; PMID:7601846
A:Accession: I40327
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-267 <RES>
A:Cross-references: UNIPROT:Q45338; UNIPARC:UPI0000038DE7; EMBL:U12020; NID:G687228; PID:
R: Allen, A.; Maskell, D.
Mol. Microbiol. 19, 37-52, 1996
A:Title: The identification, cloning and mutagenesis of a genetic locus required for lip
A:Reference number: S70669; MUID:96419162; PMID:8821935
A:Accession: S70669
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 239-267 <ALI>
A:Cross-references: UNIPARC:UPI000016E756; EMBL:X90711; NID:G992967; PIDN:CAA62242.1; PI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C:Genetics:
A:Gene: baf

Query Match 53.4%; Score 39; DB 2; Length 267;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAR 13
| | | | |
Db 195 GAIVRQWLAGR 205
| | | | |

RESULT 39
B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - *Leishmania* me
C:Species: *Leishmania mexicana*
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: B48445; S25142
R: Hannaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendoes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992
A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d
A:Reference number: A48445; MUID:93063042; PMID:1435864
A:Accession: B48445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <HAN>
A:Cross-references: UNIPROT:Q01558; UNIPARC:UPI000016BF19; EMBL:X652220; NID:G99552; PIDN
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 53.4%; Score 39; DB 2; Length 331;
Best Local Similarity 38.5%; Pred. No. 57;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAR 13
: | | | | |
Db 185 VDGPSLKDWRGGR 197
| | | | |

RESULT 40
A72514
hypothetical protein APE2086 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 05-Oct-2004
C:Accession: A72514
R: Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aeropyr*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <KAW>
A:Cross-references: UNIPROT:Q9YA52; UNIPARC:UPI000005E1E5; DBJ:AF000063; NID:G5105654;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2086
C:Superfamily: PP-loop ATPase, ydaO type

```
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82291
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82291
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-791 <HEI>
A;Cross-references: UNIPROT:Q9KU26; UNIPARC:UPI00000C2D86; GB:AE004157; GB:AE003852; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0703
A;Map position: 1

Query Match      53.4%; Score 39; DB 2; Length 791;
Best Local Similarity 63.6%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEQPTLRQWLA 11
      ||| |||
Db      730 IESETMHWQLA 740

RESULT 44
GNLJGC
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate CAM2/Gul;
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly;
C;Species: human immunodeficiency virus type 2, HIV-2
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: B38475; JQ0974
R;Tristram, M.; Hill, F.; Karpas, A.
J. Gen. Virol. 72, 721-724, 1991
A;Title: Nucleotide sequence of a Guinea-Bissau-derived human immunodeficiency virus typ;
A;Reference number: A38475; MUID:91170959; PMID:2005437
A;Accession: B38475
A;Molecule type: DNA
A;Residues: 1-1034 <TRI>
A;Cross-references: UNIPROT:P24107; UNIPARC:UPI0000131EF4
A;Note: readthrough of the terminator TGA may occur between codons ATT for 564-Ile and G;
C;Comment: The cleavage sites of this polyprotein have not been determined.
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleoc
F;85-183/Product: retropepsin #status predicted <RTP>
F;109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      53.4%; Score 39; DB 1; Length 1034;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 EGPTLRQW 9
      :||| |||
Db      200 DGPRLQW 207

RESULT 45
GNLJGG
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate GH-1)
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004
C;Accession: J50328
R;Haegawa, A.; Teujimoto, H.; Maki, N.; Ishikawa, K.; Miura, T.; Fukasawa, M.; Miki, K.
AIDS Res. Hum. Retroviruses 5, 593-604, 1989
A;Title: Sequence of a distinct HIV-2 isolate from Ghana showing significant divergence
```

```
Query Match      53.4%; Score 39; DB 2; Length 331;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy      3 GPTLRQWLAARA 14
      ||| |||
Db      306 GPPVREWFRRRA 317

RESULT 41
C87021
serine-threonine protein kinase [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87021
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S;
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: C87021
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 <STO>
A;Cross-references: UNIPROT:O69568; UNIPARC:UPI00000C60C9; GB:AL450380; NID:gl3092968; F
C;Genetics:
A;Gene: ML0897
C;Superfamily: Mycobacterium tuberculosis probable serine/threonine-specific protein kin

Query Match      53.4%; Score 39; DB 2; Length 400;
Best Local Similarity 69.2%; Pred. No. 69;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEQPTLRQWLAAR 13
      ||| |||
Db      86 IEQPTLRQWLAAR 98

RESULT 42
C83221
transport protein Haad PA3406 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 05-Oct-2004
C;Accession: C83221
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: C83221
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-600 <STO>
A;Cross-references: UNIPROT:Q9HYJ8; UNIPARC:UPI00000C59F5; GB:AE004761; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: haad; PA3406

Query Match      53.4%; Score 39; DB 2; Length 600;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEQPTLRQWLAAR 12
      :||| |||
Db      392 LDGADLRQWSAA 403

RESULT 43
A82291
c-di-GMP phosphodiesterase A-related protein VC0703 [imported] - Vibrio cholerae (strain
```

A:Reference number: J50327; MUID:90122350; PMID:2611042
A:Accession: J50328
A:Molecule type: DNA
A:Residues: 1-1035 <HAS>
A:Cross-references: UNIPROT:P18042; UNIPARC:UPI0000174A34
A:Note: This sequence was submitted to JIPID, October 1989
C:Comment: Cleavage sites that yield the mature proteins remain to be determined.
C:Genetics:
A:Gene: pol
A:Start codon: ACA
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase;
F:85-184/Product: retropepsin #status predicted <RTP>
F:109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1035;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
:|||||
Db 200 DGPRLRQW 207

RESULT 46
GNLJG2
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ROD)
N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 2, HIV-2
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C:Accession: B26262
R:Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alison, M.
Nature 326, 662-669, 1987
A:Title: Genome organization and transactivation of the human immuno-deficiency virus ty
A:Reference number: A26262; MUID:87173056; PMID:3031510
A:Accession: B26262
A:Molecule type: DNA
A:Residues: 1-1036 <GUY>
A:Cross-references: UNIPROT:P04594; UNIPARC:UPI0000131EFA; GB:M15390
C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucle
F:86-184/Product: retropepsin #status predicted <RTP>
F:110/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1036;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
:|||||
Db 201 DGPRLRQW 208

RESULT 47
GNLJST
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ST)
N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 2, HIV-2
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: B33943
R:Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;
J. Virol. 64, 890-901, 1990
A:Title: Molecular characterization of an attenuated human immunodeficiency virus type 2
A:Reference number: A33943; MUID:90112662; PMID:2296086
A:Accession: B33943
A:Molecule type: genomic RNA
A:Residues: 1-1055 <KUM>
A:Cross-references: UNIPROT:P20876; UNIPARC:UPI0000131EFC
C:Comment: The pol polyprotein contains reverse transcriptase and endonuclease. However,

C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase; revers
F:105-203/Product: retropepsin #status predicted <RTP>
F:129/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 53.4%; Score 39; DB 1; Length 1055;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
:|||||
Db 220 DGPRLRQW 227

RESULT 48
SSJ092
pol polyprotein - human immunodeficiency virus type 2
C:Species: human immunodeficiency virus type 2, HIV-2
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: SSJ092
R:Becker, M.; Zort, B.; Becker, A.; Habermehl, K.O.
submitted to the EMBL Data Library, March 1995
A:Description: Molecular and phylogenetic characterisation of a Guinea-Bissau-derived hu
A:Reference number: SSJ091
A:Accession: SSJ092
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1055 <BBC>
A:Cross-references: UNIPROT:O73194; UNIPARC:UPI00001785DA; EMBL:Z48731
C:Superfamily: pol polyprotein
C:Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 1055;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
:|||||
Db 220 DGPRLRQW 227

RESULT 49
S08436
pol polyprotein - human immunodeficiency virus type 2 D205 (fragment)
C:Species: human immunodeficiency virus type 2 D205, HIV-2 D205
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S08436
R:Dietrich, U.; Adamski, M.; Kreutz, R.; Seipp, A.; Kuehn, H.; Ruebsamen-Waigmann, H.
Nature 342, 948-950, 1989
A:Title: A highly divergent HIV-2-related isolate.
A:Reference number: S08434; MUID:90081881; PMID:2594088
A:Accession: S08436
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1058 <DIE>
A:Cross-references: UNIPROT:PI5833; UNIPARC:UPI0000131EP6; EMBL:X16109
A:Note: This sequence was submitted to the EMBL Data Library, Aug-1989
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: polyprotein

Query Match 53.4%; Score 39; DB 2; Length 1058;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAAR 13
:|||||
Db 221 DGPRLRQWPLSR 232

```

RESULT 50
T13423
hypothetical protein 30B8.4 - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C/Accession: T13423
R/Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A/Reference number: Z17668
A/Accession: T13423
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3345 <MUR>
A/Cross-references: UNIPROT:O46074; UNIPARC:UPI00000796D3; EMBL:AL009195; NID:el355203;
C/Genetics:
A/Cross-references: FlyBase:FBgn0000377
A/Map position: X
A/Introns: 51/3; 159/1; 476/1; 526/1; 1465/1; 1826/3; 1947/3; 2081/1; 2196/3; 3007/3
A/Note: EG:30B8.4

Query Match      53.4%; Score 39; DB 2; Length 3345;
Best Local Similarity 50.0%; Pred. NO. 6.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 IEQPTLRQWLAA 12
       ::|||
Db      2203 VKNPKLEQWLAS 2214

Search completed: May 12, 2006, 10:52:21
Job time : 21.0915 secs

```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds
(without alignments)

99.694 Million cell updates/sec

Title: US-10-632-388-13

Perfect score: 73

Sequence: 1 IEPTLRQWLAARA 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	68.5	53	Q5VWJ5	nocardia fa
2	50	67.8	302	Q742B3	MYCPA
3	49.5	67.8	333	CBBR_XANFL	MYCPA
4	49	67.1	319	Q9RKM5	STRCO
5	48	65.8	607	Q9L8D4	POLCB
6	48	65.8	815	Q4KS46	orange-spot
7	48	65.8	941	Q8QUJ6	9VIRU
8	47	64.4	296	Q8ZGS7	YERPE
9	47	64.4	296	Q6GD06	YERPS
10	46	63.0	306	Q7BD906	MYCTU
11	46	63.0	306	Q05576	MYCTU
12	46	63.0	306	Q7U0W3	MYCBO
13	46	63.0	580	Q89RH2	BRATA
14	45	61.6	245	Q66272	9SPHN
15	45	61.6	249	Q82989	9SPHN
16	45	61.6	278	Q9XDV0	9SPHN
17	45	61.6	421	Q7WLX1	BORPA
18	45	61.6	421	Q7WQU8	BORBP
19	45	61.6	756	Q885P2	PSBSM
20	44	60.3	207	GST3_CAEEL	CAEEL
21	44	60.3	536	ENTE_ECO57	ENTE
22	44	60.3	536	ENTE_ECOLI	ENTE
23	44	60.3	536	Q83M10	SHIGLIA fl
24	44	60.3	756	Q4ZOD5	PSBSY
25	44	60.3	760	Q7G4N4	ORYSA
26	43	58.9	285	KSQA_TREPA	TREPA
27	43	58.9	297	Q7UOE4	RHOBA
28	43	58.9	306	Q8EJ00	SHEON
29	43	58.9	354	Q8ZYT5	PYRAE
30	43	58.9	377	Q82PX5	STRAW
31	43	58.9	683	Q83436	TREPA

32	43	58.9	728	2	Q617Q3	CAEER
33	43	58.9	754	2	Q95f82	CAEEL
34	43	58.9	1157	2	Q527F0	MAGGR
35	42	57.5	91	2	Q8Y015	RALSO
36	42	57.5	126	2	Q8N9N4	HUMAN
37	42	57.5	127	2	Q4SMZ4	TETNG
38	42	57.5	252	2	Q8XPQ9	RALSO
39	42	57.5	313	2	P90433	SIVCZ
40	42	57.5	325	2	Q855N9	CAUD
41	42	57.5	326	2	P95613	RHIGA
42	42	57.5	375	2	Q7FAN4	ORYSA
43	42	57.5	375	2	Q7XPP6	ORYSA
44	42	57.5	450	2	Q9SLB9	ARATH
45	42	57.5	532	2	Q5GW75	XANOR
46	42	57.5	586	2	Q9N6P9	LEIMA
47	42	57.5	1019	1	POL	SIVS4
48	42	57.5	1019	2	P89154	SIVCZ
49	42	57.5	1019	2	Q7ZBR7	SIVCZ
50	42	57.5	1019	2	Q7ZBR5	SIVCZ
51	41.5	56.8	400	2	Q5GYS2	XANOR
52	41.5	56.8	410	2	Q4UU69	XANCP
53	41.5	56.8	410	2	Q8P915	XANCP
54	41.5	56.8	427	2	Q8PLB2	XANAC
55	41	56.2	75	2	Q98AJ1	RHILLO
56	41	56.2	89	2	Q5QJW6	ROS1
57	41	56.2	89	2	Q5QJX7	ROS1
58	41	56.2	89	2	Q5QJY2	ROS1
59	41	56.2	89	2	Q5QK18	ROS1
60	41	56.2	89	2	Q5QK33	ROS1
61	41	56.2	89	2	Q5QK53	ROS1
62	41	56.2	89	2	Q5QK54	ROS1
63	41	56.2	130	2	Q6PBK0	BRARE
64	41	56.2	136	2	Q5ZFJ3	POPTN
65	41	56.2	136	2	Q5ZFJ8	POPTN
66	41	56.2	136	2	Q5ZFJ9	POPTN
67	41	56.2	136	2	Q5ZFJ0	POPTN
68	41	56.2	136	2	Q5ZFL1	POPTN
69	41	56.2	136	2	Q5ZFL2	POPTN
70	41	56.2	136	2	Q5ZFL4	POPTN
71	41	56.2	136	2	Q5ZFM1	POPTN
72	41	56.2	136	2	Q5ZFM7	POPTN
73	41	56.2	136	2	Q5ZFM9	POPTN
74	41	56.2	136	2	Q5ZFN3	POPTN
75	41	56.2	136	2	Q5ZFN5	POPTN
76	41	56.2	136	2	Q5ZFN7	POPTN
77	41	56.2	136	2	Q5ZFN8	POPTN
78	41	56.2	137	2	Q8S7Z0	ROS1
79	41	56.2	137	2	Q8S7Z1	MANES
80	41	56.2	137	2	Q8S7Z1	MANES
81	41	56.2	137	2	Q4VJ49	9ASTR
82	41	56.2	137	2	Q4VJ48	9ASTR
83	41	56.2	137	2	Q4VJ44	9ASTR
84	41	56.2	137	2	Q4VJ41	9ASTR
85	41	56.2	137	2	Q4VJ38	9ASTR
86	41	56.2	137	2	Q4VJ35	9ASTR
87	41	56.2	137	2	Q4VJ32	9ASTR
88	41	56.2	153	2	Q80ZRO	MOUSE
89	41	56.2	160	2	Q56WJ4	ARATH
90	41	56.2	163	2	Q4FAA7	BRANA
91	41	56.2	173	2	Q8GZB8	BRARP
92	41	56.2	195	1	ACPT_ECO57	ACPT
93	41	56.2	195	1	ACPT_ECOLI	ACPT
94	41	56.2	195	1	Q8FCN3	ECOL6
95	41	56.2	195	2	Q6N108	RHOBA
96	41	56.2	207	2	Q6PKU2	OROMI
97	41	56.2	209	2	Q6N1X5	RHOBA
98	41	56.2	219	2	Q8H6A7	ORYSA
99	41	56.2	223	2	Q7XAP7	HOUCO
100	41	56.2	224	2	Q9F3Q7	STRCO
101	41	56.2	244	2	Q9R7K1	9SPHN
102	41	56.2	245	2	Q82987	9SPHN
103	41	56.2	245	2	Q82991	9SPHN
104	41	56.2	249	2	Q9A535	CAUCR

Q617Q3	caenorhabdi
Q95Y82	caenorhabdi
Q527F0	magnaporthe
Q8Y015	raistonia s
Q8N9N4	homo sapien
Q4SMZ4	tetradon n
Q8XPQ9	raistonia s
P90433	chimpanzee
Q855N9	mycobacteri
P95613	rhizobium g
Q7FAN4	oryza sativ
Q7XPP6	oryza sativ
Q9SLB9	arabidopsis
Q5GW75	xanthomonas
Q9N6P9	leishmania
P12502	simian immu
P89154	chimpanzee
Q7ZBR7	chimpanzee
Q7ZBR5	chimpanzee
Q5GYS2	xanthomonas
Q4UU69	xanthomonas
Q8P915	xanthomonas
Q8PLB2	xanthomonas
Q98AJ1	rhizobium l
Q5QJW6	shorea parv
Q5QJX7	shorea lepr
Q5QJY2	shorea lepr
Q5QK18	shorea curt
Q5QK33	shorea acum
Q5QK53	shorea acum
Q5QK54	shorea hypo
Q6PBK0	brachydanio
Q5ZFJ3	populus tre
Q5ZFJ8	populus tre
Q5ZFJ9	populus tre
Q5ZFJ0	populus tre
Q5ZFL1	populus tre
Q5ZFL2	populus tre
Q5ZFL4	populus tre
Q5ZFM1	populus tre
Q5ZFM7	populus tre
Q5ZFM9	populus tre
Q5ZFN3	populus tre
Q5ZFN5	populus tre
Q5ZFN7	populus tre
Q5ZFN8	populus tre
Q8S7Z0	manihot pru
Q8S7Z1	manihot esc
Q8S7Z1	manihot esc
Q4VJ49	scaevola ch
Q4VJ48	scaevola co
Q4VJ44	scaevola ga
Q4VJ41	scaevola ga
Q4VJ38	scaevola ki
Q4VJ35	scaevola mo
Q4VJ32	scaevola pr
Q80ZRO	mus musculu
Q56WJ4	arabidopsis
Q4FAA7	brassica ra
Q8GZB8	brassica ra
Q8X5U4	escherichia
P37623	escherichia
Q8FCN3	escherichia
Q6N108	rhodopseudo
Q6PKU2	orobanche m
Q6N1X5	rhodopseudo
Q8H6A7	oryza sativ
Q7XAP7	houttuynia
Q9F3Q7	streptomyce
Q9R7K1	erythroba
Q82987	erythroba
Q82991	erythroba
Q9A535	caulobacter

105	41	56.2	266	2	Q81D70_BACCR	Q81D70 bacillus ce	178	40	54.8	536	2	Q8FK18_ECOL6	Q8fk18 escherichia
106	41	56.2	268	2	Q7Y082_LYCES	Q7Y082 lycopersico	179	40	54.8	598	2	Q6N4Y3_RHOPA	Q6n4y3 rhodopseudo
107	41	56.2	294	2	Q04891_LYCES	Q04891 lycopersico	180	40	54.8	619	2	Q4KX0_HUMAN	Q4kx0 homo sapien
108	41	56.2	298	2	Q5YPT6_NOCFA	Q5YPT6 nocardia fa	181	40	54.8	690	1	EPB42_HUMAN	P16452 homo sapien
109	41	56.2	306	2	Q9Z5G1_MYCLE	Q9Z5G1 mycobacteri	182	40	54.8	691	2	Q4VB97_HUMAN	Q4vb97 homo sapien
110	41	56.2	308	1	XERC_CORGL	Q8nnz9 corynebacte	183	40	54.8	818	2	Q5NV45_9BURK	Q5nv45 talstonia m
111	41	56.2	313	2	Q8VMF0_CAPAN	Q8vwp0 capsicum an	184	40	54.8	882	2	Q4P0J6_USTWA	Q4p0j6 ustilago ma
112	41	56.2	326	1	G3PC_T0BAC	P09094 nicotiana t	185	40	54.8	1082	2	Q756B5_ASHGO	Q756b5 ashbaya goss
113	41	56.2	336	1	G3PC_PETCR	P26519 petroselinu	186	40	54.8	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis
114	41	56.2	336	2	Q89ZV8_BACTN	Q89zv8 bacteroides	187	40	54.8	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis
115	41	56.2	337	1	G3PC_MESCR	P17878 mesembryant	188	40	54.8	1444	2	Q4PG64_USTWA	Q4pg64 ustilago ma
116	41	56.2	337	1	G3PC_SINAL	P04796 oryza alb	189	40	54.8	1706	2	Q5VZA9_HUMAN	Q5vza9 homo sapien
117	41	56.2	337	2	G6K5G8_ORYSA	Q6k5g8 oryza sativ	190	40	54.8	1712	1	CO4A2_HUMAN	P08572 homo sapien
118	41	56.2	338	1	G3PC_ARATH	P25858 arabidopsis	191	40	54.8	1896	2	Q9DRA1_VIRU	Q9dra1 botrytis vi
119	41	56.2	338	1	G3PC_DIACA	P34921 dianthus ca	192	39.5	54.1	325	2	Q9HP97_HALSA	Q9hp97 halobacteri
120	41	56.2	338	2	Q8LAS0_ARATH	P34921 dianthus ca	193	39.5	54.1	613	2	Q89HCl_BRAJA	Q89hcl bradyrhizob
121	41	56.2	338	2	Q8LK04_SOLITU	Q8lk04 solanum tub	194	39	53.4	65	2	Q984B5_RHILO	Q984b5 rhizobium l
122	41	56.2	338	2	Q9FX54_ARATH	Q9fx54 arabidopsis	195	39	53.4	131	2	P72684_SYNY3	P72684 synechocyst
123	41	56.2	402	2	Q4NC82_9MICC	Q4nc82 arthrobacte	196	39	53.4	133	2	Q4ZPV7_PSESV	Q4zpv7 pseudomonas
124	41	56.2	485	2	Q8DSX5_VIBU	Q8dsx5 vibrio vuln	197	39	53.4	134	2	Q86G15_PSESM	Q86g15 pseudomonas
125	41	56.2	631	2	Q6LP74_PHOPR	Q6lp74 photobacter	198	39	53.4	148	2	Q4UK00_RICPE	Q4uk00 rickettsia
126	41	56.2	719	2	Q92ZH9_RHIME	Q92zh9 rhizobium m	199	39	53.4	154	2	Q9W101_9HIV2	Q9w101 human immun
127	41	56.2	750	2	Q9UG82_AGRTS	Q9ug82 agrobacteri	200	39	53.4	160	2	Q73UZ7_MYCPA	Q73uz7 mycobacteri
128	41	56.2	766	2	Q9MW20_PSEPU	Q9mw20 pseudomonas	201	39	53.4	164	2	Q6MNC9_BDEBA	Q6mnc9 bdellovibri
129	41	56.2	776	2	G6K603_ORYSA	Q6k603 oryza sativ	202	39	53.4	167	2	Q880B5_PSESM	Q880b5 pseudomonas
130	41	56.2	791	2	Q73S36_MYCPA	Q73s36 mycobacteri	203	39	53.4	168	2	Q9V492_DROME	Q9v492 drosophila
131	41	56.2	818	2	Q7WMY0_ALCEU	Q7wmy0 alcaligenes	204	39	53.4	169	2	Q4ZXM1_PSESV	Q4zxm1 pseudomonas
132	41	56.2	863	2	Q9ST50_MAIZE	Q9st50 ze mays (m	205	39	53.4	196	2	Q8SAP5_9BRYO	Q8sap5 mitthyridiu
133	41	56.2	1017	2	Q6VG40_SIVCZ	Q6vg40 chimpanzee	206	39	53.4	217	2	Q87115_SIVCZ	Q87115 chimpanzee
134	41	56.2	1153	2	Q92KF2_RHIME	Q92kf2 rhizobium m	207	39	53.4	252	2	Q8XY68_RALSO	Q8xy68 talstonia s
135	40	54.8	114	2	Q73029_BACCI	Q73029 bacillus an	208	39	53.4	259	2	Q8PF17_XANAC	Q8pf17 xanthomonas
136	40	54.8	115	2	Q81M37_BACAN	Q81m37 bacillus an	209	39	53.4	267	1	BAF_BOEPR	Q88005 bordetella
137	40	54.8	115	2	Q98HW9_RHILO	Q98hw9 rhizobium l	210	39	53.4	267	1	BAF_BOEPR	Q45338 bordetella
138	40	54.8	129	2	Q4I3M8_GIBZEL	Q4i3m8 gibberella	211	39	53.4	287	2	Q74222_MYCPA	Q74222 mycobacteri
139	40	54.8	129	2	Q8DHX7_SYNEL	Q8dhx7 synechococc	212	39	53.4	310	2	Q7SK19_9HIV2	Q7sk19 human immun
140	40	54.8	137	2	G6FE11_ACIAD	Q6fe11 acinetobact	213	39	53.4	319	2	Q89HR9_BRAJA	Q89hr9 bradyrhizob
141	40	54.8	154	2	Q54XF9_DICDI	Q54xf9 dictyosteli	214	39	53.4	326	2	Q9BVC4_9HEPC	Q9bvc4 hepatitis c
142	40	54.8	177	2	Q4UTU1_XANCP	Q4utl1 xanthomonas	215	39	53.4	330	1	G3PC_LEIME	Q01558 leishmania
143	40	54.8	177	2	Q8PA34_XANCP	Q8pa34 xanthomonas	216	39	53.4	331	2	Q9YA52_AERPE	Q9ya52 aeropyrum p
144	40	54.8	194	2	Q72J80_THET2	Q72j80 thermus the	217	39	53.4	335	2	Q82D06_STRAW	Q82d06 streptomyce
145	40	54.8	218	2	Q5RJN8_RAT	Q5rjn8 rattus norv	218	39	53.4	337	1	G3P_PHACH	Q01982 phanerocha
146	40	54.8	230	2	Q72KL1_9HIV1	Q72kl1 human immun	219	39	53.4	338	2	Q4PI23_USTWA	Q4pi23 ustilago ma
147	40	54.8	238	2	Q835J7_ENTFA	Q835j7 enterococcu	220	39	53.4	338	2	Q4LPN5_9BURK	Q4lpn5 burkholderi
148	40	54.8	239	2	Q00510_STRAM	Q00510 streptomyce	221	39	53.4	340	2	Q8UN04_SIVCZ	Q8un04 chimpanzee
149	40	54.8	241	2	Q4MI12_BACCE	Q4mi12 bacillus ce	222	39	53.4	340	2	Q8UN03_SIVCZ	Q8un03 chimpanzee
150	40	54.8	241	2	Q631B5_BACCC	Q631b5 bacillus ce	223	39	53.4	343	2	Q4LXZ8_9BURK	Q4lxz8 burkholderi
151	40	54.8	242	2	Q81K59_BACAN	Q81k59 bacillus an	224	39	53.4	349	2	Q7SKK9_9HIV2	Q7skk9 human immun
152	40	54.8	251	2	Q6HB50_BACHK	Q6hb50 bacillus th	225	39	53.4	349	2	Q7SKK6_9HIV2	Q7skk6 human immun
153	40	54.8	251	2	Q815C8_BACCR	Q815c8 bacillus ce	226	39	53.4	349	2	Q7SKK4_9HIV2	Q7skk4 human immun
154	40	54.8	271	2	Q4MVQ5_BACCE	Q4mvq5 bacillus ce	227	39	53.4	349	2	Q7SKJ6_9HIV2	Q7skj6 human immun
155	40	54.8	272	2	Q72XP5_BACCI	Q72xp5 bacillus ce	228	39	53.4	349	2	Q7SKJ5_9HIV2	Q7skj5 human immun
156	40	54.8	305	2	Q7NRM4_CHRVO	Q7nrm4 chromobacte	229	39	53.4	349	2	Q7SKJ4_9HIV2	Q7skj4 human immun
157	40	54.8	308	2	Q4NG45_9MICC	Q4ng45 arthrobacte	230	39	53.4	349	2	Q7SKJ3_9HIV2	Q7skj3 human immun
158	40	54.8	310	2	Q9D8Z7_WOUSE	Q9d8z7 mus musculu	231	39	53.4	349	2	Q7SKJ2_9HIV2	Q7skj2 human immun
159	40	54.8	331	2	Q63U03_BURPS	Q63u03 burkholderi	232	39	53.4	349	2	Q7SKJ1_9HIV2	Q7skj1 human immun
160	40	54.8	346	2	Q99M02_BRAJA	Q99m02 bradyrhizob	233	39	53.4	349	2	Q7SKJ0_9HIV2	Q7skj0 human immun
161	40	54.8	349	2	Q7SKK8_9HIV2	Q7skk8 human immun	234	39	53.4	349	2	Q7SKI8_9HIV2	Q7sk18 human immun
162	40	54.8	355	1	TSN10_HUMAN	Q9h129 homo sapien	235	39	53.4	349	2	Q7SKI7_9HIV2	Q7sk17 human immun
163	40	54.8	369	2	Q4WGC2_ASPFU	Q4wgc2 aspergillus	236	39	53.4	349	2	Q7SKI6_9HIV2	Q7sk16 human immun
164	40	54.8	369	2	Q69LD7_ORYSA	Q69ld7 oryza sativ	237	39	53.4	349	2	Q7SKI5_9HIV2	Q7sk15 human immun
165	40	54.8	398	2	Q740W3_MYCPA	Q740w3 mycobacteri	238	39	53.4	349	2	Q7SKI4_9HIV2	Q7sk14 human immun
166	40	54.8	410	2	Q4LN26_9BURK	Q4ln26 burkholderi	239	39	53.4	349	2	Q7SKI3_9HIV2	Q7sk13 human immun
167	40	54.8	410	2	Q63H26_BURPS	Q63h26 burkholderi	240	39	53.4	349	2	Q7SKI2_9HIV2	Q7sk12 human immun
168	40	54.8	410	2	Q629V1_BURMA	Q629v1 burkholderi	241	39	53.4	349	2	Q7SKI1_9HIV2	Q7sk11 human immun
169	40	54.8	411	2	Q4KDP2_PSEFS	Q4kdp2 pseudomonas	242	39	53.4	351	2	Q9E3T1_9HIV2	Q9e3t1 human immun
170	40	54.8	429	2	Q7S227_NEUCR	Q7s227 neuropep	243	39	53.4	351	2	Q9E3T2_9HIV2	Q9e3t2 human immun
171	40	54.8	441	2	Q73S55_MYCPA	Q73s55 mycobacteri	244	39	53.4	351	2	Q9E3T4_9HIV2	Q9e3t4 human immun
172	40	54.8	463	1	YKAA_EACPF	P30267 bacillus ps	245	39	53.4	353	2	O5AV13_EMENI	O5av13 aspergillus
173	40	54.8	472	2	Q9WHZ9_9HIV2	Q9whz9 human immun	246	39	53.4	353	2	Q9KMU0_GSPHN	Q9kmu0 spingomona
174	40	54.8	494	2	Q650T8_ORYSA	Q650t8 oryza sativ	247	39	53.4	354	2	Q9KZQ9_STRCO	Q9kzq9 streptomyce
175	40	54.8	530	2	Q5F558_NEIG1	Q5f558 neisseria g	248	39	53.4	374	2	Q738S3_MYCPA	Q738s3 mycobacteri
176	40	54.8	530	2	Q9JWB3_NEIMA	Q9jwb3 neisseria m	249	39	53.4	396	2	Q90PU1_SIVCZ	Q90pul chimpanzee
177	40	54.8	531	2	Q9JX19_NEIMB	Q9jx19 neisseria m	250	39	53.4	398	2	Q820L4_NITEU	Q820l4 nitrosomona

251	39	53.4	400	2	O69568 MYCLE	O69568 mycobacteri	324	39	53.4	1677	2	Q4SDD0_TETNG	Q4sdd0 tetraodon n
252	39	53.4	402	2	Q82R87_STRAW	Q82r87 streptomyc	325	39	53.4	1722	2	Q53L67_ORYSA	Q53l67 oryza sativ
253	39	53.4	403	2	Q88N20_PSEPK	Q88n20 pseudomonas	326	39	53.4	2355	2	Q4KCD0_PSEF5	Q4kcd0 pseudomonas
254	39	53.4	414	2	Q8FT01_COREF	Q8ft01 corynebacte	327	39	53.4	3433	1	PCX DROME	P18490 drosophila
255	39	53.4	457	2	Q7UL80_RHOBA	Q7ule0 rhodopirell	328	39	53.4	6481	2	Q4Q892_LEIMA	Q4qe92 leishmania
256	39	53.4	468	2	Q88H13_PSEPK	Q88h13 pseudomonas	329	38.5	52.7	535	2	Q5M2M6_SYNP6	Q5m2m6 synchococc
257	39	53.4	476	2	Q88H13_PSEPK	Q88h13 pseudomonas	330	38	52.1	61	2	Q5Y514_NOCPA	Q5y514 nocardia fa
258	39	53.4	496	2	Q6CE07_YARLI	Q6ce07 yarowia li	331	38	52.1	67	2	Q5Y514_NOCPA	Q5y514 nocardia fa
259	39	53.4	545	2	Q6N594_RHOBA	Q6n594 rhodopseu	332	38	52.1	87	2	Q8XZ44_RALSO	Q8xzf4 ralstonia s
260	39	53.4	553	2	Q7WG77_BORBR	Q7wgh7 bordetella	333	38	52.1	88	2	Q8XZ44_RALSO	Q8xzf4 ralstonia s
261	39	53.4	553	2	Q7W512_BORBA	Q7w512 bordetella	334	38	52.1	106	2	Q6H0G4_9HOMO	Q6h0g4 leccinum qu
262	39	53.4	553	2	Q7VWX5_BORPE	Q7vwx5 bordetella	335	38	52.1	106	2	Q6H0J0_9HOMO	Q6h0j0 leccinum ma
263	39	53.4	560	2	Q66347_9HIV2	Q66347 human immu	336	38	52.1	118	2	Q6H0G2_9HOMO	Q6h0g2 leccinum ca
264	39	53.4	571	2	Q7S508_NEUCR	Q7s508 neurospora	337	38	52.1	120	2	Q5I5E9_9HOMO	Q5i5e9 leccinum va
265	39	53.4	573	2	Q5FUI6_GUOX	Q5fui6 gluconobact	338	38	52.1	123	2	Q5I5F7_9HOMO	Q5i5f7 leccinum sc
266	39	53.4	579	2	Q6WLD7_BACT	Q6wld7 uncultured	339	38	52.1	124	2	Q6H0H8_9HOMO	Q6h0h8 leccinum ro
267	39	53.4	582	2	Q8AB47_BACTN	Q8ab47 bacteroides	340	38	52.1	129	2	Q6H0G0_9HOMO	Q6h0g0 leccinum mo
268	39	53.4	586	2	Q828W4_STRAW	Q828w4 streptomyc	341	38	52.1	129	2	Q6I360_DROME	Q6i360 drosophila
269	39	53.4	600	2	Q9HYJ8_PSEAE	Q9hyj8 pseudomonas	342	38	52.1	132	2	Q8NIQ4_HUMAN	Q8nlq4 homo sapien
270	39	53.4	609	2	Q856X8_9CAUD	Q856x8 mycobacteri	343	38	52.1	132	2	Q5I5G9_9HOMO	Q5i5g9 leccinum sc
271	39	53.4	609	2	Q889R7_RHILO	Q889r7 rhizobium l	344	38	52.1	133	2	Q944T3_FRAAN	Q944t3 fragaria an
272	39	53.4	612	2	Q830U7_ENTFA	Q830j7 enterococcu	345	38	52.1	133	2	Q944T7_FRAAN	Q944t7 fragaria an
273	39	53.4	618	2	Q4LQV6_9BURK	Q4lqv6 burkholderi	346	38	52.1	133	2	Q944T7_FRAAN	Q944t7 fragaria an
274	39	53.4	632	2	Q75I19_ORYSA	Q75i19 oryza sativ	347	38	52.1	134	2	Q944T7_FRAAN	Q944t7 fragaria an
275	39	53.4	647	2	Q6FV06_CANGA	Q6fv06 candida gla	348	38	52.1	140	2	Q5I5H2_9HOMO	Q5i5h2 leccinum ro
276	39	53.4	693	2	Q5TNY2_ANOGA	Q5tny2 anopheles g	349	38	52.1	140	2	Q6H0F8_9HOMO	Q6h0f8 leccinum in
277	39	53.4	706	2	Q63U33_BURPS	Q63u33 burkholderi	350	38	52.1	141	2	Q5I5F4_9HOMO	Q5i5f4 leccinum sc
278	39	53.4	706	2	Q62KAI_BURMA	Q62kai burkholderi	351	38	52.1	141	2	Q5I5F1_9HOMO	Q5i5f1 leccinum va
279	39	53.4	713	2	Q7NL87_GLOVI	Q7nlr0 gloeobacter	352	38	52.1	142	2	Q7XWC2_ORYSA	Q7xwc2 oryza sativ
280	39	53.4	725	2	Q8WR00_LEIMA	Q8wr00 leishmania	353	38	52.1	144	2	Q6H0G3_9HOMO	Q6h0g3 leccinum au
281	39	53.4	733	2	Q5XQF4_MAIZE	Q5xqf4 zea mays (m	354	38	52.1	146	2	Q6L9U0_9ENTR	Q6l9u0 bremeria q
282	39	53.4	740	2	Q8XVE2_RALSO	Q8xve2 ralstonia s	355	38	52.1	149	2	Q81924_CICAR	Q81924 cicar arlet
283	39	53.4	766	2	Q8XVE2_RALSO	Q8xve2 ralstonia s	356	38	52.1	194	2	BETI_PSEAE	Q88af0 pseudomonas
284	39	53.4	766	2	Q8XVE2_RALSO	Q8xve2 ralstonia s	357	38	52.1	197	1	BETI_PSEAE	Q88af0 pseudomonas
285	39	53.4	766	2	Q8XVE2_RALSO	Q8xve2 ralstonia s	358	38	52.1	197	1	BETI_PSEAE	Q88af0 pseudomonas
286	39	53.4	788	2	Q89QX5_BRAJA	Q89qx5 bradyrhizob	359	38	52.1	197	1	BETI_PSEAE	Q88af0 pseudomonas
287	39	53.4	791	2	Q8KU26_VIBCH	Q8ku26 vibrio chol	360	38	52.1	197	1	BETI_PSEAE	Q88af0 pseudomonas
288	39	53.4	925	2	Q7XSI0_ORYSA	Q7xsi0 oryza sativ	361	38	52.1	214	2	Q6H0H9_9HOMO	Q6h0h9 leccinum ve
289	39	53.4	986	2	Q57059_SIVCZ	Q57059 chimpanzee	362	38	52.1	214	2	Q6H0H9_9HOMO	Q6h0h9 leccinum ve
290	39	53.4	1005	2	Q6Y8X5_9HIV1	Q6y8x5 human immu	363	38	52.1	218	2	Q6H0H1_9HOMO	Q6h0h1 leccinum pa
291	39	53.4	1015	2	Q8JAH1_SIVCZ	Q8jah1 chimpanzee	364	38	52.1	225	2	Q6H0H1_9HOMO	Q6h0h1 leccinum pa
292	39	53.4	1022	1	POL_SIVSP	P19505 simian immu	365	38	52.1	225	2	Q6H0H1_9HOMO	Q6h0h1 leccinum pa
293	39	53.4	1022	2	Q90317_SIVCZ	Q90317 chimpanzee	366	38	52.1	227	2	Q9RBZ4_PSESG	Q9rbz4 pseudomonas
294	39	53.4	1022	2	Q87956_SIVCZ	Q87956 chimpanzee	367	38	52.1	227	2	Q9RBZ4_PSESG	Q9rbz4 pseudomonas
295	39	53.4	1022	2	Q87965_SIVCZ	Q87965 chimpanzee	368	38	52.1	227	2	Q9RBZ4_PSESG	Q9rbz4 pseudomonas
296	39	53.4	1022	2	Q89620_SIVCZ	Q89620 chimpanzee	369	38	52.1	227	2	Q9RBZ4_PSESG	Q9rbz4 pseudomonas
297	39	53.4	1022	2	Q88135_SIVCZ	Q88135 chimpanzee	370	38	52.1	231	2	Q88BV4_PSESM	Q88bv4 pseudomonas
298	39	53.4	1034	1	POL_HV2CA	P24107 human immu	371	38	52.1	231	2	Q88BV4_PSESM	Q88bv4 pseudomonas
299	39	53.4	1035	2	Q73194_9HIV2	Q73194 human immu	372	38	52.1	231	2	Q88BV4_PSESM	Q88bv4 pseudomonas
300	39	53.4	1035	2	Q6R793_9HIV2	Q6r793 human immu	373	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
301	39	53.4	1035	2	Q6R784_9HIV2	Q6r784 human immu	374	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
302	39	53.4	1036	1	POL_HV2RO	P04584 human immu	375	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
303	39	53.4	1049	1	POL_HV2G1	P18042 human immu	376	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
304	39	53.4	1049	1	POL_HV2G1	P18042 human immu	377	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
305	39	53.4	1055	1	POL_HV2ST	P20876 human immu	378	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
306	39	53.4	1055	2	Q4YU00_9HIV2	Q4yuo0 human immu	379	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
307	39	53.4	1056	2	Q40097_SIVCZ	Q40097 chimpanzee	380	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
308	39	53.4	1058	1	POL_HV2D2	P15833 human immu	381	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
309	39	53.4	1059	2	Q76630_9HIV2	Q76630 human immu	382	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
310	39	53.4	1060	2	Q5QG01_SIVCZ	Q5qg01 chimpanzee	383	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
311	39	53.4	1060	2	Q5QPT5_SIVCZ	Q5qpt5 chimpanzee	384	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
312	39	53.4	1060	2	Q5QFL2_SIVCZ	Q5qfl2 chimpanzee	385	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
313	39	53.4	1060	2	Q5QFJ6_SIVCZ	Q5qfj6 chimpanzee	386	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
314	39	53.4	1060	2	Q5QFJ6_SIVCZ	Q5qfj6 chimpanzee	387	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
315	39	53.4	1060	2	Q70NFB_PHYPA	Q70nfb physcomitre	388	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
316	39	53.4	1060	2	Q4WNFB_ASPFU	Q4wnfb aspergillus	389	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
317	39	53.4	1209	2	Q6H0K6_9HIV2	Q6h0k6 human immu	390	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
318	39	53.4	1226	1	NPHN MOUSE	Q9qzs7 mus musculu	391	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
319	39	53.4	1256	2	Q9JIX1_MOUSE	Q9jix1 mus musculu	392	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
320	39	53.4	1414	2	Q7S131_NEUCR	Q7s131 neurospora	393	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
321	39	53.4	1431	2	Q5B491_EMENI	Q5b491 aspergillus	394	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
322	39	53.4	1435	2	Q4WIA7_ASPFU	Q4wia7 aspergillus	395	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas
323	39	53.4	1447	2	Q4FWA4_LEIMA	Q4fwa4 leishmania	396	38	52.1	233	2	Q88BV4_PSESM	Q88bv4 pseudomonas

397	38	52.1	268	2	Q5I5G5_9HOMO	Q5I5G5 leccinum sc	470	38	52.1	414	2	Q4L061_9HIV1	Q4L061 human immun
398	38	52.1	273	2	Q5I5G1_9HOMO	Q5I5G1 leccinum sc	471	38	52.1	423	2	Q8AH22_9HIV1	Q8AH22 human immun
399	38	52.1	274	2	Q6H012_9HOMO	Q6H012 leccinum ve	472	38	52.1	433	2	Q4IWE5_AZOV1	Q4IWE5 azotobacteri
400	38	52.1	276	2	Q6H014_9HOMO	Q6H014 leccinum qu	473	38	52.1	436	2	Q73R19_MYCPA	Q73R19 mycobacteri
401	38	52.1	277	2	Q6H0J3_9HOMO	Q6H0J3 leccinum du	474	38	52.1	449	2	Q5P209_AZOSE	Q5P209 azoarcus sp
402	38	52.1	276	2	Q4NYJ6_9DEL1	Q4NYJ6 anaeromyxob	475	38	52.1	452	2	Q7NMK8_GLOVI	Q7NMK8 gloeobacter
403	38	52.1	278	2	Q4F8T9_9LECA	Q4F8T9 heterodea m	476	38	52.1	462	2	Q4LYF0_9BURK	Q4LYF0 burkholderi
404	38	52.1	280	2	Q6H0H0_9HOMO	Q6H0H0 leccinum sc	477	38	52.1	467	2	Q5Z206_NOCAPA	Q5Z206 nocardia fa
405	38	52.1	281	2	Q5I5I2_9HOMO	Q5I5I2 leccinum br	478	38	52.1	469	2	Q8Y92_9ANASP	Q8Y92 anabaena sp
406	38	52.1	281	2	Q6H015_9HOMO	Q6H015 leccinum au	479	38	52.1	476	2	Q4WL94_ASPPU	Q4WL94 aspergillus
407	38	52.1	281	2	Q6H016_9HOMO	Q6H016 leccinum pi	480	38	52.1	478	2	Q7US57_RHOBA	Q7US57 rhodopirell
408	38	52.1	282	2	Q5I5B6_9HOMO	Q5I5B6 leccinum va	481	38	52.1	491	2	Q7U9L8_SYNEPC	Q7U9L8 synechococc
409	38	52.1	282	2	Q5I5H6_9HOMO	Q5I5H6 leccinum ho	482	38	52.1	493	2	Q4WYB6_ASPPU	Q4WYB6 aspergillus
410	38	52.1	282	2	Q6H017_9HOMO	Q6H017 leccinum vu	483	38	52.1	494	2	Q9ESC6_MOUSE	Q9ESC6 mus musculu
411	38	52.1	282	2	Q6H019_9HOMO	Q6H019 leccinum ar	484	38	52.1	496	2	Q846W9_STRCM	Q846W9 streptomyce
412	38	52.1	283	2	Q5I5H9_9HOMO	Q5I5H9 leccinum ho	485	38	52.1	519	2	Q8A7A4_BACTN	Q8A7A4 bacteroides
413	38	52.1	283	2	Q6H0H4_9HOMO	Q6H0H4 leccinum br	486	38	52.1	550	2	Q73YZ0_MYCPA	Q73YZ0 mycobacteri
414	38	52.1	285	2	Q6H0H2_9HOMO	Q6H0H2 leccinum ho	487	38	52.1	567	2	Q4T902_9SPHN	Q4T902 erythrobract
415	38	52.1	287	2	Q6H018_9HOMO	Q6H018 leccinum vu	488	38	52.1	577	2	Q4HCU3_9DEIO	Q4HCU3 deinococcus
416	38	52.1	295	2	Q04I06_LYCES	Q04I06 lycopersico	489	38	52.1	580	2	Q4SYV1_TETNG	Q4SYV1 tetraodon n
417	38	52.1	295	2	Q5VAL5_PANGI	Q5VAL5 panax ginse	490	38	52.1	588	2	Q53QB5_ORYSA	Q53QB5 oryza sativ
418	38	52.1	297	1	XERC_MYCLE	Q9CBU0 mycobacteri	491	38	52.1	589	2	Q9A402_CAUCR	Q9A402 caulobacter
419	38	52.1	299	2	Q8Y305_RALSO	Q8Y305 ralstonia s	492	38	52.1	595	2	Q4FTB6_9GAMM	Q4FTB6 psychobact
420	38	52.1	300	2	Q5RPD1_9HIV1	Q5RPD1 human immun	493	38	52.1	600	2	Q8ITJ3_HETGL	Q8ITJ3 heterodera
421	38	52.1	301	2	Q73V05_MYCPA	Q73V05 mycobacteri	494	38	52.1	621	2	Q7VZ40_BORPE	Q7VZ40 bordetella
422	38	52.1	304	2	Q4LP46_9BURK	Q4LP46 burkholderi	495	38	52.1	621	2	Q7W8T9_BORBA	Q7W8T9 bordetella
423	38	52.1	305	1	XERD_RHILO	Q98IX8 rhizobium l	496	38	52.1	621	2	Q7WK49_BORBR	Q7WK49 bordetella
424	38	52.1	309	2	Q602I8_METCA	Q602I8 methylococc	497	38	52.1	631	2	Q5FT72_GLUOX	Q5FT72 gluconobact
425	38	52.1	311	1	LYSR_ECOLI	P03030 escherichia	498	38	52.1	635	2	Q9A882_CAUCR	Q9A882 caulobacter
426	38	52.1	311	2	Q7AB47_ECO57	Q7AB47 escherichia	499	38	52.1	636	2	Q595V7_9VIRU	Q595V7 pleurotous o
427	38	52.1	311	2	Q57KA3_SALCH	Q57KA3 salmonella	500	38	52.1	644	2	Q4FXG0_LEIMA	Q4FXG0 leishmania
428	38	52.1	311	2	Q8X6J2_ECO57	Q8X6J2 escherichia	501	38	52.1	750	2	Q7UUG1_RHOBA	Q7UUG1 rhodopirell
429	38	52.1	311	2	Q8FEA1_ECOL6	Q8FEA1 escherichia	502	38	52.1	764	2	Q4WK40_ASPPU	Q4WK40 aspergillus
430	38	52.1	311	2	Q88B21_PSESM	Q88B21 pseudomonas	503	38	52.1	816	2	Q59068_PYROCO	Q59068 pyrococcus
431	38	52.1	311	2	Q7N7G0_PROHL	Q7N7G0 photorhabdu	504	38	52.1	817	2	Q98KX1_RHILO	Q98KX1 rhizobium l
432	38	52.1	311	2	Q7CPV5_SALTY	Q7CPV5 salmonella	505	38	52.1	869	2	Q9VZ82_DROME	Q9VZ82 drosophila
433	38	52.1	311	2	Q5PEP7_SALOPA	Q5PEP7 salmonella	506	38	52.1	880	2	Q6MSK4_CORGL	Q6MSK4 corynebacte
434	38	52.1	311	2	Q8XGD5_SALTI	Q8XGD5 salmonella	507	38	52.1	904	1	DPO1_MYCBO	DPO1 mycobacteri
435	38	52.1	311	2	Q83QB6_SHIGLA	Q83QB6 shigella fl	508	38	52.1	904	1	DPO1_MYCTU	DPO1 mycobacteri
436	38	52.1	313	2	Q667E7_YERSIN	Q667E7 yersinia ps	509	38	52.1	905	2	Q8NQS1_CORGL	Q8NQS1 corynebacte
437	38	52.1	314	1	Y2599_MYCTU	P0A595 mycobacteri	510	38	52.1	926	2	Q9ATQ3_WHEAT	Q9ATQ3 triticum ae
438	38	52.1	314	1	Q84BW1_BURML	P0A596 mycobacteri	511	38	52.1	933	2	Q677U1_LYPHOCY	Q677U1 lymphocyati
439	38	52.1	315	2	Q5FPV4_GLUOX	Q5FPV4 burkholderi	512	38	52.1	955	2	Q7V7J0_PROMM	Q7V7J0 prochloroco
440	38	52.1	316	2	Q8VWPI_CAPSIC	Q8VWPI gluconobact	513	38	52.1	962	2	Q6U676_HUMAN	Q6U676 homo sapien
441	38	52.1	317	2	Q8VWPI_CAPSIC	Q8VWPI capsicum an	514	38	52.1	998	2	Q8JBD2_9HIV1	Q8JBD2 human immun
442	38	52.1	319	2	Q6N2R8_RHOPE	Q6N2R8 rhodopseudo	515	38	52.1	1018	2	Q9IBL9_9HIV2	Q9IBL9 human immun
443	38	52.1	328	2	Q8CZK8_YERPE	Q8CZK8 yersinia pe	516	38	52.1	1057	1	POL_SIVAI	POL_SIVAI simian immu
444	38	52.1	329	2	Q9DCZ1_9HIV1	Q9DCZ1 human immun	517	38	52.1	1102	2	Q7YZ34_CRYPV	Q7YZ34 cryptospori
445	38	52.1	329	2	Q9DCZ2_9HIV1	Q9DCZ2 human immun	518	38	52.1	1155	2	P94183_9NOST	P94183 anabaena sp
446	38	52.1	331	2	Q6NJ86_CORDI	Q6NJ86 corynebacte	519	38	52.1	1155	2	Q8YMH0_ANASP	Q8YMH0 anabaena sp
447	38	52.1	332	2	Q7ZMA0_9HIV1	Q7ZMA0 human immun	520	38	52.1	1156	2	Q5CL84_CRYHO	Q5CL84 cryptospori
448	38	52.1	334	2	Q5QZ31_IDILO	Q5QZ31 idiomarina	521	38	52.1	1176	2	Q5CJW8_CRYPV	Q5CJW8 cryptospori
449	38	52.1	335	2	Q4JDJ2_9PROT	Q4JDJ2 nitrosomona	522	38	52.1	1208	2	Q5B104_EMENI	Q5B104 aspergillus
450	38	52.1	336	2	Q7MT61_PORGI	Q7MT61 porphyromon	523	38	52.1	1234	1	NPHN_RAT	NPHN rattus norv
451	38	52.1	337	1	G3PC_ANTMA	P25861 antirrhinum	524	38	52.1	1235	2	Q52FK0_MAGGR	Q52FK0 magnaporthe
452	38	52.1	337	1	G3PC_MAIZE	P08735 zea mays (m	525	38	52.1	1241	2	Q9HFV7_COLGL	Q9HFV7 colletotric
453	38	52.1	337	1	G3PC_PETHY	P6520 petunia hyb	526	38	52.1	1252	2	Q9JIX2_RAT	Q9JIX2 rattus norv
454	38	52.1	337	1	G3PD_MAIZE	Q9054 zea mays (m	527	38	52.1	1256	2	Q9Z5S5_MOUSE	Q9Z5S5 mus musculu
455	38	52.1	337	1	G3P_USTMA	P09317 ustilago ma	528	38	52.1	1266	2	Q9ET59_MOUSE	Q9ET59 mus musculu
456	38	52.1	337	2	Q4PBW2_USTMA	Q4PBW2 ustilago ma	529	38	52.1	1263	2	Q5KZP8_GEOKA	Q5KZP8 geobacillus
457	38	52.1	337	2	Q6RUO2_DAUCA	Q6RUO2 daucus caro	530	38	52.1	1287	2	Q4NRJ4_9DELT	Q4NRJ4 anaeromyxob
458	38	52.1	337	2	Q8XG67_TOBAC	Q8XG67 nicotiana t	531	38	52.1	1304	2	Q7SDK1_NEUCR	Q7SDK1 neurospora
459	38	52.1	339	2	Q9IILH_PSEAE	Q9IILH pseudomonas	532	38	52.1	1319	2	Q61EX3_CAEBR	Q61EX3 caenorhabdi
460	38	52.1	350	1	MUTY_ECOLI	P17802 escherichia	533	38	52.1	1391	2	Q5AVS1_EMENI	Q5AVS1 aspergillus
461	38	52.1	350	2	Q8XC58_ECO57	Q8XC58 escherichia	534	38	52.1	1578	2	Q4ZUT5_PSESY	Q4ZUT5 pseudomonas
462	38	52.1	352	1	ID12_PYRAE	Q8ZYF6 pyrobaculum	535	38	52.1	3781	2	Q9VIS5_DROME	Q9VIS5 drosophila
463	38	52.1	352	2	Q67DX1_9RALS	Q67DX1 ralstonia s	536	38	52.1	4268	2	Q8XVP2_RALSO	Q8XVP2 ralstonia s
464	38	52.1	360	1	G3P_ATRNU	Q34783 atriplex nu	537	37.5	51.4	93	2	Q71R12_BACSU	Q71R12 bacillus su
465	38	52.1	360	2	Q8FE21_ECOL6	Q8FE21 escherichia	538	37.5	51.4	294	2	Q4HCZ0_9DEIO	Q4HCZ0 deinococcus
466	38	52.1	392	2	Q6W157_RHISN	Q6W157 rhizobium s	539	37.5	51.4	311	2	Q4H9C0_9DEIO	Q4H9C0 deinococcus
467	38	52.1	392	2	Q8PP76_XANAC	Q8PP76 xanthomonas	540	37.5	51.4	436	2	Q45293_CORGL	Q45293 corynebacte
468	38	52.1	399	2	Q6SF36_9BACT	Q6SF36 uncultured	541	37.5	51.4	482	2	Q7X9B3_MEDTR	Q7X9B3 medicago tr
469	38	52.1	403	2	Q5DN43_9CAUD	Q5DN43 bacterioph	542	37.5	51.4	533	1	OPGD_XANAC	OPGD xanthomonas

543	37.5	51.4	555	2	Q7QHJ7 ANOGA	Q7ghj7 anopheles g	616	37	50.7	233	2	Q9DQ2_9H1V1	Q9dqr2 human immun
544	37	50.7	79	2	Q53301_9ROSI	Q533q1 quercus gla	617	37	50.7	234	2	Q4K3V9_PSEFS	Q4k3v9 pseudomonas
545	37	50.7	79	2	Q533R2_9ROSI	Q533r2 quercus gla	618	37	50.7	235	2	Q9M4V4_WHEAT	Q9m4v4 triticum ae
546	37	50.7	81	2	Q3NDL7_HYDMA	Q3ndl7 hydra magni	619	37	50.7	235	1	Q4J4K4_AZOVI	Q4j4k4 azobacter
547	37	50.7	85	2	Q64GN7_9H1V1	Q64gn7 human immun	620	37	50.7	237	1	PYRF_GLOVI	P7nk22 gloebacter
548	37	50.7	87	2	Q65930_CAEV	Q65930 caprine art	621	37	50.7	239	2	Q5MI13_9H1V1	Q5mi13 human immun
549	37	50.7	87	2	Q9WR47_9RETR	Q9wr47 brazilian c	622	37	50.7	240	2	Q5MI07_9H1V1	Q5mi07 human immun
550	37	50.7	103	2	Q95QV6_CAEEL	Q95qv6 caenorhabdi	623	37	50.7	243	2	Q4ZFX5_PSESY	Q4zpx5 pseudomonas
551	37	50.7	110	2	Q856U7_9CAUD	Q856u7 mycobacteri	624	37	50.7	243	2	Q75808_9H1V1	Q75808 human immun
552	37	50.7	110	2	Q8H246_VACMY	Q8h246 vaccinium m	625	37	50.7	243	2	Q75801_9H1V1	Q75801 human immun
553	37	50.7	111	2	Q86Z67_HETAN	Q86z67 heterobasid	626	37	50.7	244	2	Q4V8S1_BRARE	Q4v8s1 brachydanio
554	37	50.7	111	2	Q86ZQ4_HETAN	Q86zq4 heterobasid	627	37	50.7	245	1	KDSB_FUSNN	Q8rfa8 fusobacteri
555	37	50.7	111	2	Q86ZQ5_HETAN	Q86zq5 heterobasid	628	37	50.7	245	2	Q6H0J6_9HOMO	Q6h0j6 leccinum cr
556	37	50.7	111	2	Q86ZQ6_HETAN	Q86zq6 heterobasid	629	37	50.7	245	2	Q6V776_HORSP	Q6v776 hordeum spo
557	37	50.7	120	2	Q506S1_9ROSA	Q506s1 ficus popen	630	37	50.7	245	2	Q6V785_HORSP	Q6v785 hordeum spo
558	37	50.7	120	2	Q506T1_9ROSA	Q506t1 ficus popen	631	37	50.7	245	2	Q7P7V1_FUSNV	Q7p7v1 fusobacteri
559	37	50.7	120	2	Q506U4_9ROSA	Q506u4 ficus cf. t	632	37	50.7	245	2	Q6Q7J9_9H1V1	Q6qtj9 human immun
560	37	50.7	120	2	Q506U5_9ROSA	Q506u5 ficus inslp	633	37	50.7	245	2	Q58S25_9H1V1	Q58s25 human immun
561	37	50.7	120	2	Q506U6_9ROSA	Q506u6 ficus glabr	634	37	50.7	245	2	Q75785_9H1V1	Q75785 human immun
562	37	50.7	120	2	Q506U7_9ROSA	Q506u7 ficus yopon	635	37	50.7	247	2	Q75807_9H1V1	Q75807 human immun
563	37	50.7	120	2	Q506U8_9ROSA	Q506u8 ficus maxim	636	37	50.7	247	2	Q6EFI6_9H1V1	Q6efi6 human immun
564	37	50.7	120	2	Q506U9_9ROSA	Q506u9 ficus dugan	637	37	50.7	248	2	Q6H0J5_9HOMO	Q6h0j5 leccinum ca
565	37	50.7	120	2	Q506V0_9ROSA	Q506v0 ficus trigo	638	37	50.7	249	2	Q903C2_9H1V1	Q903c2 human immun
566	37	50.7	120	2	Q506V1_9ROSA	Q506v1 ficus citri	639	37	50.7	250	2	Q5MHX9_9H1V1	Q5mhx9 human immun
567	37	50.7	120	2	Q506V2_9ROSA	Q506v2 ficus parae	640	37	50.7	253	2	Q7ZKF3_9H1V1	Q7zkf3 human immun
568	37	50.7	120	2	Q506V3_9ROSA	Q506v3 ficus colub	641	37	50.7	255	2	Q90LK9_9H1V1	Q90lk9 human immun
569	37	50.7	120	2	Q506V4_9ROSA	Q506v4 ficus costa	642	37	50.7	256	2	Q56922_9H1V1	Q56922 human immun
570	37	50.7	120	2	Q506V5_9ROSA	Q506v5 ficus perfo	643	37	50.7	259	2	Q5RRV0_9H1V1	Q5rrv0 human immun
571	37	50.7	120	2	Q506V6_9ROSA	Q506v6 ficus pertu	644	37	50.7	261	2	Q5LVM4_SILPO	Q5lrv4 silicibacte
572	37	50.7	120	2	Q506V7_9ROSA	Q506v7 ficus obtus	645	37	50.7	263	2	Q6EFL4_9H1V1	Q6efl4 human immun
573	37	50.7	127	2	Q97401_9H1V1	Q97401 human immun	646	37	50.7	266	2	Q98LGI_RHILO	Q98lgi rhizobium l
574	37	50.7	127	2	Q97396_9H1V1	Q97396 human immun	647	37	50.7	266	2	Q4SAB1_TETNG	Q4sa81 tetraodon l
575	37	50.7	128	2	Q6G0X8_BAROU	Q6g0x8 bartonella	648	37	50.7	270	2	Q5GYG8_XANOR	Q5gyg8 xanthomonas
576	37	50.7	133	2	Q944T8_FRAAN	Q944t8 fragaria an	649	37	50.7	272	2	Q3E1S2_DESVH	Q3e1s2 lupinus alb
577	37	50.7	133	2	Q91Y39_RAT	Q91y39 rattus norv	650	37	50.7	292	2	Q53I52_LUPAL	Q5i52 lupinus alb
578	37	50.7	135	2	Q52FK0_POPTN	Q52fk0 populus tre	651	37	50.7	295	2	Q9EPG8_9H1V1	Q9epg8 human immun
579	37	50.7	136	2	Q82BS3_STRAW	Q82bs3 streptomyce	652	37	50.7	295	2	Q4IBQ0_GIBZE	Q4ibq0 gibberella
580	37	50.7	138	2	Q9GSP8_HYDAT	Q9gsp8 hydra atten	653	37	50.7	296	2	Q03307_CHELO	Q03307 chelodina l
581	37	50.7	139	2	Q9L2H5_STRCO	Q9l2h5 streptomyce	654	37	50.7	297	2	Q03335_PELUSIO	Q03335 pelusios wi
582	37	50.7	139	2	Q9L2H5_STRCO	Q9l2h5 streptomyce	655	37	50.7	297	2	Q03335_PELUSIO	Q03335 pelusios wi
583	37	50.7	140	2	Q515F3_9HOMO	Q515f3 leccinum sn	656	37	50.7	298	2	Q8J800_9H1V1	Q8j800 human immun
584	37	50.7	140	2	Q82BN7_STRAW	Q82bn7 streptomyce	657	37	50.7	299	2	Q5RMX3_9H1V1	Q5rmx3 human immun
585	37	50.7	142	2	Q6VZA0_ORYSA	Q6vza0 oryza sativ	658	37	50.7	300	2	Q6R664_9H1V1	Q6r664 human immun
586	37	50.7	142	2	Q6VZA0_ORYSA	Q6vza0 oryza sativ	659	37	50.7	300	2	Q5RRL4_9H1V1	Q5rrl4 human immun
587	37	50.7	147	2	Q82YB1_XANCP	Q82yb1 xanthomonas	660	37	50.7	300	2	Q5RRL4_9H1V1	Q5rrl4 human immun
588	37	50.7	147	2	Q82YB1_XANCP	Q82yb1 xanthomonas	661	37	50.7	301	2	Q4NHT3_9M1CC	Q4nht3 arthrobacte
589	37	50.7	147	2	Q82YB1_XANCP	Q82yb1 xanthomonas	662	37	50.7	302	2	Q8KLC3_RH1ET	Q8klc3 rhizobium e
590	37	50.7	151	2	Q72880_9H1V1	Q72880 human immun	663	37	50.7	302	2	Q8KLC3_RH1ET	Q8klc3 rhizobium e
591	37	50.7	152	2	Q809F3_9H1V1	Q809f3 human immun	664	37	50.7	303	2	Q8J821_9H1V1	Q8j821 human immun
592	37	50.7	152	2	Q809F3_9H1V1	Q809f3 human immun	665	37	50.7	304	2	Q5RRZ7_9H1V1	Q5rrz7 human immun
593	37	50.7	165	2	Q9A5T0_CAUCR	Q9a5t0 caulobacter	666	37	50.7	304	2	Q5RRZ7_9H1V1	Q5rrz7 human immun
594	37	50.7	166	2	Q6TIB6_9H1V1	Q6tib6 human immun	667	37	50.7	305	1	G3PC_HORVU	G3pc hordeum vul
595	37	50.7	168	2	Q7XJJI_WHEAT	Q7xjji triticum ae	668	37	50.7	308	2	Q3NQ8_BURPS	Q3nq8 burkholderi
596	37	50.7	168	2	Q9B596_CHILA	Q9b596 chinchilla	669	37	50.7	308	2	Q6ZAA0_BURMA	Q6zaa0 burkholderi
597	37	50.7	168	2	Q4TFP2_TETNG	Q4tfp2 tetraodon n	670	37	50.7	308	2	Q6R6F9_9H1V1	Q6r6f9 human immun
598	37	50.7	173	2	Q8PLV3_XANAC	Q8plv3 xanthomonas	671	37	50.7	313	2	Q8VNZ1_TREHY	Q8vnz1 treponema h
599	37	50.7	173	2	Q8PLV3_XANAC	Q8plv3 xanthomonas	672	37	50.7	315	2	Q8VM39_9GP1R	Q8vm39 brachyspira
600	37	50.7	175	2	Q69MK4_ORYSA	Q69mk4 oryza sativ	673	37	50.7	315	2	Q98ZB3_9H1V1	Q98zb3 human immun
601	37	50.7	179	2	Q886G5_PSESM	Q886g5 pseudomonas	674	37	50.7	318	2	Q6WEW0_9STRA	Q6we0 apodachlya
602	37	50.7	182	2	Q9B598_CHILA	Q9b598 chinchilla	675	37	50.7	319	2	Q76EG1_9EUGL	Q76eg1 eutrepriell
603	37	50.7	190	2	Q9SHP3_ARATH	Q9shp3 arabidopsis	676	37	50.7	319	2	Q50A19_9H1V1	Q50a19 human immun
604	37	50.7	192	2	Q7NT62_CHRVO	Q7nt62 chromobacte	677	37	50.7	320	2	Q8BYH8_9H1V1	Q8byh8 human immun
605	37	50.7	195	2	Q955I9_9NEOB	Q955i9 indirana sp	678	37	50.7	321	2	Q8JB46_9H1V1	Q8jb46 human immun
606	37	50.7	195	2	Q6QUAI_COCIM	Q6qjai coccidoide	679	37	50.7	321	2	Q8Q60_9H1V1	Q8q60 human immun
607	37	50.7	197	2	Q83J79_SHIFL	Q83j79 shigella fl	680	37	50.7	323	2	Q8XQW6_RALSO	Q8xqw6 ralstonia s
608	37	50.7	203	2	Q76EK6_9LAMI	Q76ek6 torenia hyb	681	37	50.7	324	2	Q6WPY7_9H1V1	Q6wp7 human immun
609	37	50.7	206	2	Q622A8_CAEER	Q622a8 caenorhabdi	682	37	50.7	328	2	Q9ED10_9H1V1	Q9ed10 human immun
610	37	50.7	207	2	Q7XHP1_ARYAD	Q7xhp1 oryza sativ	683	37	50.7	328	2	Q9ECQ1_9H1V1	Q9ecq1 human immun
611	37	50.7	217	2	Q6FEV9_ACIAD	Q6fev9 acinetobact	684	37	50.7	328	2	Q9EE13_9H1V1	Q9ee13 human immun
612	37	50.7	222	2	Q5RM10_9H1V1	Q5rm10 human immun	685	37	50.7	328	2	Q9EIN1_9H1V1	Q9ein1 human immun
613	37	50.7	226	2	Q8Q3M7_9H1V1	Q8q3m7 human immun	686	37	50.7	328	2	Q9EFH1_9H1V1	Q9efh1 human immun
614	37	50.7	226	2	Q8Q3N4_9H1V1	Q8q3n4 human immun	687	37	50.7	328	2	Q9EE76_9H1V1	Q9ee76 human immun
615	37	50.7	229	2	Q9WK80_9H1V1	Q9wk80 human immun	688	37	50.7	328	2	Q9EEW3_9H1V1	Q9ee76 human immun

689	37	50.7	328	2	Q9ECX4_9H1V1	Q9ecx4 human immun	762	37	50.7	384	2	Q9J3Z1_9H1V1	Q9j3z1 human immun
690	37	50.7	328	2	Q9EC5_9H1V1	Q9ec5 human immun	763	37	50.7	386	1	ETR1_CANTR	Q8wzm3 candida tro
691	37	50.7	328	2	Q9EKU1_9H1V1	Q9ek1 human immun	764	37	50.7	386	1	ETR2_CANTR	Q8wzm4 candida tro
692	37	50.7	328	2	Q9ECU5_9H1V1	Q9ecj5 human immun	765	37	50.7	387	2	Q90GP0_9H1V1	Q90gp0 human immun
693	37	50.7	328	2	Q9EHX6_9H1V1	Q9ehx6 human immun	766	37	50.7	391	2	Q9S2B3_STRCO	Q9s2b3 streptomyce
694	37	50.7	328	2	Q9EL92_9H1V1	Q9el92 human immun	767	37	50.7	391	2	Q98Y28_9H1V1	Q98y28 human immun
695	37	50.7	328	2	Q9EDJ2_9H1V1	Q9edj2 human immun	768	37	50.7	392	2	Q36575_9H1V1	Q36575 human immun
696	37	50.7	328	2	Q9EH6_9H1V1	Q9ehc6 human immun	769	37	50.7	392	2	Q36575_METHA	Q36575 methanoearc
697	37	50.7	328	2	Q9ECF3_9H1V1	Q9ecf3 human immun	770	37	50.7	395	2	Q5FQ25_GLOOX	Q5fq25 gluconobact
698	37	50.7	328	2	Q9EGH3_9H1V1	Q9egh3 human immun	771	37	50.7	398	2	Q4FDM6_DROSI	Q4fdm6 xerophilu
699	37	50.7	328	2	Q904D4_9H1V1	Q904d4 human immun	772	37	50.7	398	2	Q5M7R8_XENTR	Q5m7r8 xenopus tro
700	37	50.7	329	2	Q9QB7_9H1V1	Q9qbr7 human immun	773	37	50.7	401	2	Q7VVG2_BORPE	Q7vvg2 bordetella
701	37	50.7	330	2	Q50AS4_9H1V1	Q50as4 human immun	774	37	50.7	401	2	Q7WB96_BORPA	Q7wb96 bordetella
702	37	50.7	331	1	G3P_ASHGO	Q75712 ashbya goss	775	37	50.7	401	2	Q7WBR4_BORER	Q7wbr4 bordetella
703	37	50.7	331	2	Q5PY03_MUSAC	Q5py03 musa acumin	776	37	50.7	401	2	Q90IW3_9H1V1	Q90iw3 human immun
704	37	50.7	331	2	Q4JU09_CORJK	Q4ju09 corynebacte	777	37	50.7	405	2	Q5LQ49_SILPO	Q5lq49 silicibacte
705	37	50.7	332	2	Q4WDX7_ASPFU	Q4wdx7 aspergillus	778	37	50.7	406	2	Q4HCE3_9DEIO	Q4hce3 deinococcu
706	37	50.7	332	2	Q8L6S7_9UGL	Q8l6s7 euglena sp.	779	37	50.7	407	2	Q9SA29_ARATH	Q9sa29 arabidopsai
707	37	50.7	332	2	Q8Q9H6_9H1V1	Q8q9h6 human immun	780	37	50.7	409	2	Q90MT8_9H1V1	Q90mt8 human immun
708	37	50.7	332	2	Q8Q9J8_9H1V1	Q8q9j8 human immun	781	37	50.7	410	2	Q4LLO6_9BURK	Q4ll06 burkholderi
709	37	50.7	332	2	Q8Q9H4_9H1V1	Q8q9h4 human immun	782	37	50.7	411	2	Q6H703_ORYSA	Q6h703 oryza sativ
710	37	50.7	332	2	Q8Q9H3_9H1V1	Q8q9h3 human immun	783	37	50.7	413	2	Q6LGY3_PHOPR	Q6lgy3 photobacter
711	37	50.7	332	2	Q8Q9I9_9H1V1	Q8q9i9 human immun	784	37	50.7	414	2	Q87FL6_VIBPA	Q87fl6 vibrio para
712	37	50.7	332	2	Q8Q9I2_9H1V1	Q8q9i2 human immun	785	37	50.7	415	2	Q655W2_ORYSA	Q655w2 oryza sativ
713	37	50.7	332	2	Q7ZM79_9H1V1	Q7zm79 human immun	786	37	50.7	416	2	Q8DHK5_SYNEL	Q8dhk5 synechococc
714	37	50.7	332	2	Q7ZM65_9H1V1	Q7zm65 human immun	787	37	50.7	417	2	Q8AEC2_9H1V1	Q8aec2 human immun
715	37	50.7	332	2	Q7ZM00_9H1V1	Q7zm00 human immun	788	37	50.7	419	2	Q8VWN9_CAPAN	Q8vwn9 capsicum an
716	37	50.7	332	2	Q7ZL28_9H1V1	Q7zl28 human immun	789	37	50.7	420	2	Q5E924_ARATH	Q5e924 arabidopsai
717	37	50.7	332	2	Q7ZLX1_9H1V1	Q7zlx1 human immun	790	37	50.7	420	2	Q84WR0_ARATH	Q84wr0 arabidopsai
718	37	50.7	332	2	Q9FVP0_ASCNO	Q9fvp0 ascomyllum	791	37	50.7	420	2	Q8L8W7_ARATH	Q8l8w7 arabidopsai
719	37	50.7	333	2	Q8JA16_9H1V1	Q8ja16 human immun	792	37	50.7	422	2	Q9SAJ6_ARATH	Q9saj6 arabidopsai
720	37	50.7	335	1	G3P1_TRINO	P17729 trichoderma	793	37	50.7	423	2	Q9QCV1_9H1V1	Q9qcv1 human immun
721	37	50.7	335	2	Q8QAQ1_9H1V1	Q8qaq1 human immun	794	37	50.7	423	2	Q8AHP1_9H1V1	Q8ahp1 human immun
722	37	50.7	335	2	Q8QAP6_9H1V1	Q8gap6 human immun	795	37	50.7	423	2	Q8AH66_9H1V1	Q8ah66 human immun
723	37	50.7	335	2	Q8QAO8_9H1V1	Q8qa08 human immun	796	37	50.7	425	2	Q8AH61_9H1V1	Q8ah61 human immun
724	37	50.7	335	2	Q533H4_9H1V1	Q533h4 human immun	797	37	50.7	426	2	Q854V2_3CAUD	Q854v2 mycobacteri
725	37	50.7	337	1	G3PC_CRAPL	Q42671 craterostig	798	37	50.7	426	2	Q4V257_9H1V1	Q4v257 human immun
726	37	50.7	337	1	G3PC_ORYSA	Q42977 oryza sativ	799	37	50.7	433	2	Q37264_PINSY	Q37264 pinus sylve
727	37	50.7	337	1	G3PE_MAIZE	Q43247 sea mayas (m	800	37	50.7	433	2	Q37265_PINSY	Q37265 pinus sylve
728	37	50.7	337	1	G3PX_HORVU	P26517 hordeum vul	801	37	50.7	434	2	Q4JK01_9H1V1	Q4jk01 human immun
729	37	50.7	337	2	Q43359_MAIZE	Q43359 sea mayas (m	802	37	50.7	434	2	Q5D836_9H1V1	Q5d836 human immun
730	37	50.7	337	2	Q6ZK60_ORYSA	Q6zk60 oryza sativ	803	37	50.7	438	2	Q9AA92_CAUCR	Q9aa92 caulobacter
731	37	50.7	337	2	Q7FAN2_ORYSA	Q7fah2 oryza sativ	804	37	50.7	443	2	Q4UR06_XANCP	Q4ur06 xanthomonas
732	37	50.7	338	1	G3PC_PEA	P34922 pisum sativ	805	37	50.7	443	2	Q8P330_XANCP	Q8p330 xanthomonas
733	37	50.7	338	1	G3PC_RANAC	P26521 ranunculus	806	37	50.7	452	2	Q3YH86_CHICK	Q3yh86 gallus gall
734	37	50.7	340	2	Q50BB8_9H1V1	Q50bb8 human immun	807	37	50.7	460	2	Q93P09_PROMP	Q93p09 prochloroco
735	37	50.7	341	1	G3PC_CHLRE	P49644 chlamydomon	808	37	50.7	461	2	Q60CR2_METCA	Q60cr2 methylococc
736	37	50.7	341	1	G3PC_MAGLI	P26518 magnolia li	809	37	50.7	465	2	Q7V6H9_PROMM	Q7v6h9 prochloroco
737	37	50.7	341	2	Q8ZFT1_YERPE	Q8zft1 yersinia pe	810	37	50.7	477	2	Q88GH3_PSEPK	Q88gh3 pseudomonas
738	37	50.7	341	2	Q869L7_YERPS	Q869l7 yersinia ps	811	37	50.7	480	2	Q22813_ARATH	Q22813 arabidopsai
739	37	50.7	341	2	Q7NNT5_GLOVI	Q7nnt5 gloeobacter	812	37	50.7	481	2	Q8L117_RHOSH	Q8l117 rhodobacter
740	37	50.7	344	2	Q7VKH5_HAEDU	Q7vkh5 haemophilus	813	37	50.7	486	2	Q4MM98_ASPFU	Q4mm98 aspergillus
741	37	50.7	345	2	Q6DA67_BRWCT	Q6da67 erwinia car	814	37	50.7	486	2	Q9SGQ0_ARATH	Q9sgq0 arabidopsai
742	37	50.7	349	2	Q5R0S2_IDILO	Q5r0s2 idiomarina	815	37	50.7	486	2	Q5H1W5_XANOR	Q5h1w5 xanthomonas
743	37	50.7	349	2	Q90M47_9H1V1	Q90m47 human immun	816	37	50.7	491	2	Q7MC22_VIBVU	Q7mc22 vibrio vuln
744	37	50.7	351	2	Q98ZF3_9H1V1	Q98zf3 human immun	817	37	50.7	491	2	Q5L085_GEOKA	Q5l085 geobacillus
745	37	50.7	351	2	Q8Q9K2_9H1V1	Q8q9k2 human immun	818	37	50.7	499	2	Q8RT41_9H1V1	Q8rt41 human immun
746	37	50.7	352	2	Q43111_EUGGR	Q43111 euglena gra	819	37	50.7	499	2	Q67D18_9H1V1	Q67d18 human immun
747	37	50.7	353	2	Q33163_PSEAE	Q33163 pseudomonas	820	37	50.7	499	2	Q67D37_9H1V1	Q67d37 human immun
748	37	50.7	356	2	Q8AE6_9H1V1	Q8ae6 human immun	821	37	50.7	499	2	Q67D81_9H1V1	Q67d81 human immun
749	37	50.7	360	2	Q90KA3_9H1V1	Q90ka3 human immun	822	37	50.7	499	2	Q67CP5_9H1V1	Q67cp5 human immun
750	37	50.7	362	1	Q5GG09_9H1V1	Q5gg09 human immun	823	37	50.7	499	2	Q67CB1_9H1V1	Q67cb1 human immun
751	37	50.7	363	1	MRAY_AZOSE	Q5P623 azoarcus sp	824	37	50.7	499	2	Q67CA2_9H1V1	Q67ca2 human immun
752	37	50.7	363	2	Q8D0Q7_YERPE	Q8d0q7 yersinia pe	825	37	50.7	499	2	Q511W4_9H1V1	Q511w4 human immun
753	37	50.7	364	1	MRAW_BORPA	Q7wfr5 bordetella	826	37	50.7	505	2	Q51IT3_9H1V1	Q51it3 human immun
754	37	50.7	364	1	MRAW_BORER	Q7w4a7 bordetella	827	37	50.7	505	2	Q7UGI8_RHOBA	Q7ugi8 rhodopirell
755	37	50.7	364	1	MRAW_BORPE	Q7vup6 bordetella	828	37	50.7	511	2	Q83C84_COXBU	Q83c84 coxiella bu
756	37	50.7	365	2	Q65876_PINSY	Q65876 pinus sylve	829	37	50.7	513	2	Q7T2M9_CARAU	Q7t2m9 caraseius a
757	37	50.7	366	2	Q6LNZ5_PROPR	Q6lnz5 photobacter	830	37	50.7	524	2	Q9IDU5_9H1V1	Q9id5 leishmania
758	37	50.7	375	2	Q6QW16_AZOBR	Q6qw16 azospirillum	831	37	50.7	531	2	Q4Q9J5_LEIMA	Q4q9j5 burkholderi
759	37	50.7	380	2	Q7U4M1_SYNPX	Q7u4m1 synechococc	832	37	50.7	531	2	Q63V30_BURPS	Q63v30 burkholderi
760	37	50.7	382	2	Q8U337_PYRFU	Q8u337 pyrococcus	833	37	50.7	531	2	Q62JL9_BURMA	Q62jl9 burkholderi
761	37	50.7	383	2	Q53NA8_ORYSA	Q53na8 oryza sativ	834	37	50.7	547	2	Q5X8S7_STRCO	Q5x8s7 streptomyce

835	37	50.7	552	2	Q7NX80	CHRVO	Q7nx80	chromobacte	908	37	50.7	1043	2	Q90DD2	SIVCZ	Q90dd2	chimpanzee
836	37	50.7	556	2	Q6ONX3	CABR	Q6onx9	caenorhabdi	909	37	50.7	1055	2	Q89928	9HIV2	Q89928	human immun
837	37	50.7	559	2	Q8UTW0	9HIV1	Q8utw0	human immun	910	37	50.7	1061	2	Q76634	9HIV2	Q76634	human immun
838	37	50.7	560	2	Q5DNK6	9HIV1	Q5dnk6	human immun	911	37	50.7	1064	2	Q8UTW7	9HIV2	Q8utw7	human immun
839	37	50.7	562	2	Q99B10	9HIV1	Q99b10	human immun	912	37	50.7	1086	2	Q7UIK4	RHOBA	Q7uik4	rhodopirell
840	37	50.7	564	2	Q9J4V0	9HIV1	Q9j4v0	human immun	913	37	50.7	1091	1	RHG07	RAT	Q63744	r rho-stpas
841	37	50.7	569	2	Q9WA84	ARATH	Q9wa84	arabidopsis	914	37	50.7	1116	2	Q4V0E3	XANCP	Q4v0e3	xanthomonas
842	37	50.7	574	1	SYE	AREPE	Q9y9h1	aeropyrum p	915	37	50.7	1116	2	Q8PE49	XANCP	Q8pe49	xanthomonas
843	37	50.7	577	2	Q9IN28	9HIV1	Q9in28	human immun	916	37	50.7	1123	2	Q66933	9RETR	Q66933	feline immu
844	37	50.7	586	2	Q70ER5	9HIV1	Q70er5	human immun	917	37	50.7	1168	2	Q7SHS2	NEUCR	Q7shs2	neurospora
845	37	50.7	596	2	Q9M7R8	ACHBI	Q9m7r8	achlya bise	918	37	50.7	1170	2	Q4INJ1	GIBZE	Q4injl	gibberella
846	37	50.7	597	2	Q98K08	RHIL0	Q98k08	rhizobium l	919	37	50.7	1233	1	VLI	REOVD	P15024	reovirus ty
847	37	50.7	614	2	Q9LE93	PHATR	Q9le93	phaeodactyl	920	37	50.7	1275	2	Q9WAB0	REOVD	Q9wab0	reovirus ty
848	37	50.7	615	2	Q9M7R5	ODOSI	Q9m7r5	odontella s	921	37	50.7	1275	2	Q9WAB2	REOVL	Q9wab2	reovirus ty
849	37	50.7	619	2	Q4FZ41	LEIMA	Q4fz41	leishmania	922	37	50.7	1275	2	Q9WAB1	REOVJ	Q9wab1	reovirus ty
850	37	50.7	620	2	Q4FZ41	LEIMA	Q4fz41	leishmania	923	37	50.7	1357	2	Q9WAB1	REOVJ	Q9wab1	reovirus ty
851	37	50.7	630	2	Q5AVJ2	EMENI	Q5avj2	aspergillus	924	37	50.7	1391	2	Q9WAB2	REOVL	Q9wab2	reovirus ty
852	37	50.7	657	2	Q6JSB8	9HIV1	Q6jsb8	human immun	925	37	50.7	1413	2	Q4UJ10	THEAN	Q4uj10	theileria p
853	37	50.7	657	2	Q6JSB7	9HIV1	Q6jsb7	human immun	926	37	50.7	1413	2	Q4UJ10	THEAN	Q4uj10	theileria p
854	37	50.7	657	2	Q6JSB6	9HIV1	Q6jsb6	human immun	927	37	50.7	1435	2	Q4UJ10	THEAN	Q4uj10	theileria p
855	37	50.7	657	2	Q6JSB5	9HIV1	Q6jsb5	human immun	928	37	50.7	1663	2	Q88L55	PSEPK	Q88l55	pseudomonas
856	37	50.7	657	2	Q6JSB4	9HIV1	Q6jsb4	human immun	929	37	50.7	1955	2	Q4PE01	USTMA	Q4pe01	ustilago ma
857	37	50.7	657	2	Q6JSA7	9HIV1	Q6jsa7	human immun	930	37	50.7	2108	2	Q7VEZ3	MYCBO	Q7vez3	mycobacteri
858	37	50.7	657	2	Q6JSA6	9HIV1	Q6jsa6	human immun	931	37	50.7	2108	2	Q53901	MYCTU	Q53901	mycobacteri
859	37	50.7	664	2	Q9K9Z0	BACHD	Q9k9z0	bacillus ha	932	37	50.7	2229	2	Q9EMP5	STRCO	Q9emp5	streptomyce
860	37	50.7	669	2	Q526R7	MAGGR	Q526r7	magnaporthe	933	37	50.7	2307	2	Q612L3	CABBR	Q612l3	caenorhabdi
861	37	50.7	679	2	Q51U94	MAGGR	Q51u94	magnaporthe	934	37	50.7	2329	1	PRP8	CABER	P34369	caenorhabdi
862	37	50.7	700	2	Q7Q9W9	ANOGA	Q7q9w9	anopheles g	935	37	50.7	2363	2	Q4187	SCHPO	O14187	schizosacch
863	37	50.7	706	2	Q7Q9W3	HUMAN	Q7q9w3	homo sapien	936	37	50.7	3165	2	Q5AR63	EMENI	Q5ar63	aspergillus
864	37	50.7	710	2	Q8RY11	ARATH	Q8ry11	arabidopsis	937	36.5	50.0	88	2	Q6YS15	ORYSA	Q6ys15	oryza sativ
865	37	50.7	712	2	Q7MU14	PORGI	Q7mu14	porphyromon	938	36.5	50.0	133	2	Q92RI6	RHIME	Q92ri6	rhizobium m
866	37	50.7	715	2	Q9U6D9	DROME	Q9u6d9	drosophila	939	36.5	50.0	305	1	UVSE	DEIRA	Q9rte6	deinococcus
867	37	50.7	715	2	Q9U6D9	DROME	Q9u6d9	drosophila	940	36.5	50.0	339	2	Q4J2R2	AZOVI	Q4j2r2	azotobacter
868	37	50.7	715	2	Q9U6D9	DROME	Q9u6d9	drosophila	941	36.5	50.0	423	2	Q4LQN8	9BURK	Q4lqn8	burkholderi
869	37	50.7	724	2	Q5CPL7	CRYPV	Q5cpl7	cryptospori	942	36.5	50.0	529	2	Q9A4M5	CAUCR	Q9a4m5	caulobacter
870	37	50.7	724	2	Q5CK52	CRVHO	Q5ck52	cryptospori	943	36.5	50.0	529	2	Q8PQZ7	XANAC	Q8pqz7	xanthomonas
871	37	50.7	730	2	Q4PBC2	USTMA	Q4pbc2	ustilago ma	944	36.5	50.0	2221	2	Q5AV42	EMENI	Q5av42	aspergillus
872	37	50.7	749	2	Q5FR28	GLUOX	Q5fr28	gluconobact	945	36	49.3	60	1	RL28	BACST	P23374	bacillus st
873	37	50.7	755	2	Q5SR90	CRYNE	Q5sr90	cryptococcu	946	36	49.3	61	1	RL28	LECOA	O510r4	geobacillus
874	37	50.7	755	2	Q5KEY0	CRYNE	Q5key0	cryptococcu	947	36	49.3	61	1	RL28	LECOA	Q74in6	lactobacill
875	37	50.7	762	2	Q912P9	PSEAE	Q912p9	pseudomonas	948	36	49.3	77	2	Q8HZ44	PAPHA	Q8hz44	papio hamad
876	37	50.7	770	2	Q881Z6	PSESM	Q881z6	pseudomonas	949	36	49.3	81	2	Q9NDL5	9CNID	Q9ndl5	tima formos
877	37	50.7	777	1	METE	CAUCR	Q9aaw1	caulobacter	950	36	49.3	81	2	Q9NDL8	HYDEC	Q9ndl8	hydractinia
878	37	50.7	785	2	Q7NPF6	GLOVI	Q7npf6	gloeobacter	951	36	49.3	81	2	Q9NDL9	9CNID	Q9ndl9	eiylene sp.
879	37	50.7	804	2	Q4FC05	USTMA	Q4fc05	ustilago ma	952	36	49.3	85	2	Q64GR5	9HIV1	Q64gr5	human immun
880	37	50.7	825	1	5E5	RAT	Q63003	rattus norv	953	36	49.3	93	2	Y2256	XANCP	Q64gr5	human immun
881	37	50.7	829	2	Q4RX45	TETNG	Q4rx45	tetraodon n	954	36	49.3	93	2	Q4UVK1	XANCP	Q4uvk1	xanthomonas
882	37	50.7	838	2	Q7MRT1	WOLSU	Q7mrt1	wolinella s	955	36	49.3	95	1	Y2359	XANAC	Q8pvk1	xanthomonas
883	37	50.7	853	2	Q8Q5Q4	9RETR	Q8q5q4	feline immu	956	36	49.3	97	2	Q8PFC4	COREF	Q8pfc4	corynebacte
884	37	50.7	854	2	Q6J4Y9	9RETR	Q6j4y9	feline immu	957	36	49.3	98	2	Q59U7	9BACT	Q59u7	uncultured
885	37	50.7	858	2	Q9T0D8	9ALPH	Q9t0d8	gallid herp	958	36	49.3	102	2	Q3J137	MYCLE	O33137	mycobacteri
886	37	50.7	858	2	Q9T0D8	9ALPH	Q9t0d8	gallid herp	959	36	49.3	104	2	Q92U58	RHIME	Q92u58	rhizobium m
887	37	50.7	890	2	Q4URT7	XANCP	Q4urt7	xanthomonas	960	36	49.3	117	2	Q82C58	STRAW	Q82c58	streptomyce
888	37	50.7	890	2	Q7U6X4	SYNPK	Q7u6x4	synecchococ	961	36	49.3	119	2	Q8N4D3	HUMAN	Q8n4d3	homo sapien
889	37	50.7	890	2	Q8PBR5	XANCP	Q8pbr5	xanthomonas	962	36	49.3	120	2	Q6ZSJ1	HUMAN	Q6zsj1	homo sapien
890	37	50.7	907	2	Q9G6Z2	9HIV1	Q9g6z2	human immun	963	36	49.3	121	2	Q5V089	HALMA	Q5v089	haloarcula
891	37	50.7	912	2	Q9G6Z2	9HIV1	Q9g6z2	human immun	964	36	49.3	126	2	Q7QQ70	GLALA	Q7qq70	giardia lam
892	37	50.7	911	2	Q7WY21	PSEAE	Q7wy21	pseudomonas	965	36	49.3	127	2	Q6G5H7	BARHE	Q6g5h7	bartonella
893	37	50.7	940	2	Q7UQT9	RHOBA	Q7uqt9	rhodopirell	966	36	49.3	128	1	NB5M	HUMAN	Q95168	homo sapien
894	37	50.7	956	1	DPOL	DEIRA	P52027	deinococcus	967	36	49.3	129	2	Q5R6P3	PONPY	Q5r6p3	pongo pygma
895	37	50.7	999	2	Q8UTG9	9HIV1	Q8utg9	human immun	968	36	49.3	131	2	Q9CVZ3	MOUSE	Q9cvz3	mus musculu
896	37	50.7	1000	2	Q6S7Y7	9HIV1	Q6s7y7	human immun	969	36	49.3	132	2	Q6NPB9	EBEBA	Q6npb9	bdellovibri
897	37	50.7	1002	2	Q4QX95	9HIV1	Q4qx95	human immun	970	36	49.3	133	1	EPPI	HUMAN	Q95d11	macaca mula
898	37	50.7	1002	2	Q9QX95	9HIV1	Q9qx95	human immun	971	36	49.3	133	1	EPPI	HUMAN	Q95d11	macaca mula
899	37	50.7	1003	2	Q6PR23	9HIV1	Q6pr23	human immun	972	36	49.3	133	2	Q8HZ45	PAPPA	Q8hz45	papio papio
900	37	50.7	1003	2	Q519G4	9HIV1	Q519g4	human immun	973	36	49.3	134	2	Q5LF77	BACFN	Q5lf77	bacteroides
901	37	50.7	1006	2	Q5U8G6	9HIV1	Q5u8g6	human immun	974	36	49.3	135	2	Q8TN98	METAC	Q8tn98	methanosarc
902	37	50.7	1006	2	Q596M2	9HIV1	Q596m2	human immun	975	36	49.3	135	2	Q6FSL6	CANGA	Q6fsl6	candida gia
903	37	50.7	1007	2	Q66QF6	9HIV1	Q66qf6	human immun	976	36	49.3	142	2	Q8CY09	ECOL6	Q8cy09	escherichia
904	37	50.7	1009	1	POL	SIVGB	P23382	simian immu	977	36	49.3	152	2	Q82X11	NITEUP	Q82x11	nitrosomona
905	37	50.7	1009	2	Q9NKU7	LEIMA	Q9nku7	leishmania	978	36	49.3	153	2	Q9CAH9	ARATH	Q9cah9	arabidopsis
906	37	50.7	1016	2	Q8JAI0	SIVCZ	Q8jai0	chimpanzee	979	36	49.3	155	2	Q4IF93	GIBZE	Q4if93	gibberella
907	37	50.7	1039	2	Q87108	SIVCZ	Q87108	chimpanzee	980	36	49.3	156	2	Q76KB7	NYMAL	Q76kb7	nymphaea al

```
981 36 49.3 163 2 Q749N7 GEOSL
982 36 49.3 172 2 Q67FF4 _RHIV1
983 36 49.3 176 2 Q56A5 _ORISA
984 36 49.3 176 2 Q6SG12 _BACT
985 36 49.3 176 2 Q6SH12 _BACT
986 36 49.3 179 2 Q6XSL7 _RALSO
987 36 49.3 183 2 Q62KY8 _BURMA
988 36 49.3 185 2 Q8CME3 _SHEW
989 36 49.3 186 2 Q703F2 _LINUS
990 36 49.3 188 2 Q9PH16 _XYLFA
991 36 49.3 189 2 Q8CZ94 _STRR6
992 36 49.3 189 2 Q97T07 _STRPN
993 36 49.3 193 2 Q7NVL3 _CHRVO
994 36 49.3 193 2 Q4VTU6 _9SAUR
995 36 49.3 194 2 Q5SJF1 _THET8
996 36 49.3 197 2 Q2NS0 _RHIME
997 36 49.3 203 2 Q4VTU5 _9SAUR
998 36 49.3 204 1 THIE_RH1ET
999 36 49.3 204 2 Q8VLJ0 _RALSO
1000 36 49.3 207 2 Q61YTL _CAEBR
```

ALIGNMENTS

```
RESULT 1
Q5YJW5_NOCFA PRELIMINARY; PRT; 53 AA.
AC Q5YJW5;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=nfa25990;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD57446.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 53 AA; 5806 MW; 47B08B1F9BA4FCE6 CRC64;
```

```
Query Match 68.5%; Score 50; DB 2; Length 53;
Best Local Similarity 64.3%; Pred. No. 0.68;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 IEQPTLRQWLAAARA 14
: ||| |||||
Db 31 VHGPQCRLWLAARA 44
```

```
RESULT 2
Q742B3_MYCPA PRELIMINARY; PRT; 302 AA.
AC Q742B3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE GalU.
GN Name=gallU; OrderedLocusNames=MAP0924;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
```

```
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=k10;
RC Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RA SUBMITTED (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE017230; AA03241.1; -; Genomic DNA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR GO; GO:0005835; NTP_transferase.
DR InterPro; IPR005835; NTP_transferase; 1.
DR Pfam; PF00483; NTP_transferase; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;

Query Match 68.5%; Score 50; DB 2; Length 302;
Best Local Similarity 81.8%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAAR 13
||| |||||
Db 286 GPDLRQWLVAR 296

RESULT 3
CBBR_XANFL STANDARD; PRT; 333 AA.
AC P25545;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE HTH-type transcriptional regulator cbbR (RuBisCO operon
DE transcriptional regulator).
GN Name=cbbR; Synonyms=cfxO;
OS Xanthobacter flavus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Xanthobacteraceae; Xanthobacter.
OX NCBI_TaxID=281;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H4-14;
RX MEDLINE=94012468; PubMed=8407781;
RA van den Bergh E., Dijkhuizen L., Meijer W.G.;
RT "CbbR, a lysR-type transcriptional activator, is required for
RT expression of the autotrophic CO2 fixation enzymes of Xanthobacter
RT flavus.";
RL J. Bacteriol. 175:6097-6104(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-150.
RC STRAIN=H4-14;
RX MEDLINE=91172133; PubMed=1900916; DOI=10.1007/BF00269865;
RA Meijer W.G., Arnberg A.C., Enequist H.G., Terpetra P., Lidstrom M.E.,
RA Dijkhuizen L.;
RT "Identification and organization of carbon dioxide fixation genes in
RT Xanthobacter flavus H4-14.";
RL Mol. Gen. Genet. 225:320-330(1991).
CC -!- FUNCTION: Transcriptional activator for the cbb operon (cbbLSXFP)
CC for RuBisCO and other Calvin cycle genes. Binds specifically to
CC two binding sites in the cbbR-cbbL intergenic region.
CC -!- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; Z22705; CAA80406.1; -; Genomic DNA.
DR EMBL; X17252; -; NOT_ANNOTATED_CDS; Genomic DNA.
DR PIR; A36925; A36925.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
```

RESULT 5

GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 296 AA; 31378 MW; 45947413DCD54CFF CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred.No.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PTLQWLAARA 14
| | | | | | | |
Db 66 PTLQWAAASA 76

RESULT 9
Q66D06.YERPS
ID Q66D06_YERPS PRELIMINARY; PRT; 296 AA.
AC Q66D06;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative drug/metabolite (DME family) efflux pump precursor.
GN OrderedLocusNames=ypr1243;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Lazimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chantal-Francoise V., Souza B., Dacheux D., Elliott J.M.,
RA Dierise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RL comparison with Yersinia pseudotuberculosis.";
EMBL; BX336398; CAH20493.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 25 Potential.
SQ SEQUENCE 296 AA; 31407 MW; 4D3E486D32DBAC11 CRC64;

Query Match 64.4%; Score 47; DB 2; Length 296;
Best Local Similarity 81.8%; Pred.No.13;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PTLQWLAARA 14
| | | | | | | |
Db 66 PTLQWAAASA 76

RESULT 10
Q7D906.MYCTU
ID Q7D906_MYCTU PRELIMINARY; PRT; 306 AA.
AC Q7D906;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9).
GN Name=gall; OrderedLocusNames=WT1022;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK45269.1; -; Genomic_DNA.
DR TIGR; MT1022; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0003983; F:UTP-glucose-1-phosphate uridylyltransferase. . .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32406 MW; 880D3BB88CB0A3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAAR 13
DB 290 GPDLLRWLVAR 300

RESULT 11
ID O05576 MYCTU PRELIMINARY; PRT; 306 AA.
AC O05576
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE PROBABLE UTP--GLUCOSE-1-PHOSPHATE URIDYLTRANSFERASE GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE)
DE (EC 2.7.7.9).
GN Name=gau; OrderedLocusNames=Rv0993;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holtz S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842575; CAB08153.1; -; Genomic_DNA.
DR FIR; D70601; D70601.
DR TubercuList; Rv0993; -.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase.
DR Pfam; PF00483; NTP transferase; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32378 MW; 24C2387443B0A3E8 CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAAR 13
DB 290 GPDLLRWLVAR 300

RESULT 12
ID Q7UOW3 MYCBO PRELIMINARY; PRT; 306 AA.
AC Q7UOW3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE UTP--GLUCOSE-1-PHOSPHATE URIDYLTRANSFERASE GALU (UDP-
DE GLUCOSE PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
DE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PYROPHOSPHORYLASE)
DE (EC 2.7.7.9).
GN Name=gau; OrderedLocusNames=Mb1020;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthey S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248337; CAD93881.1; -; Genomic_DNA.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005835; NTP transferase; 1.
DR Pfam; PF00483; NTP transferase; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 306 AA; 32406 MW; 880D3BB88CB0A3EA CRC64;

Query Match 63.0%; Score 46; DB 2; Length 306;
Best Local Similarity 72.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 GPTLRQWLAAAR 13
DB 290 GPDLLRWLVAR 300

RESULT 13
ID Q89RH2 BRAJA PRELIMINARY; PRT; 580 AA.
AC Q89RH2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bll2800 protein.
GN OrderedLocusNames=bll12800;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RX Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
SQ SEQUENCE 580 AA; 32378 MW; 24C2387443B0A3E8 CRC64;
```

```
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
DR EMBL; BA000040; BAC48065.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008266; Tyr_Pkinase_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00481; PP2C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Transferase.
SQ SEQUENCE 580 AA; 64917 MW; 6AD3A06E8FAE143B CRC64;

Query Match 63.0%; Score 46; DB 2; Length 580;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWL 10
Db 355 IEQTTLRQWM 364

RESULT 14
O66272 9SPHN
ID O66272 9SPHN PRELIMINARY; PRT; 245 AA.
AC O66272;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14332;
RA MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB010981; BAA25791.1; -; Genomic_DNA.
DR HSP; P02954; 1QOV.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 249 AA; 27702 MW; 4D68EDC82B7166AD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 249;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWL 10
Db 26 IEQPTLNPWL 35

RESULT 16
Q9XDV0 9SPHN
ID Q9XDV0 9SPHN PRELIMINARY; PRT; 278 AA.
AC Q9XDV0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit.
GN Name=pufL;
OS Erythrobacter sp. MBIC3960.
```

```
FT NON_TER 1
SQ SEQUENCE 245 AA; 27214 MW; 52B268713E199ABD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 245;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWL 10
Db 26 IEQPTLNPWL 35

RESULT 15
O82989 9SPHN
ID O82989 9SPHN PRELIMINARY; PRT; 249 AA.
AC O82989;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3019;
RA MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB015708; BAA32995.1; -; Genomic_DNA.
DR HSP; P02954; 1YST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 249 AA; 27702 MW; 4D68EDC82B7166AD CRC64;

Query Match 61.6%; Score 45; DB 2; Length 249;
Best Local Similarity 80.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWL 10
Db 26 IEQPTLNPWL 35

RESULT 16
Q9XDV0 9SPHN
ID Q9XDV0 9SPHN PRELIMINARY; PRT; 278 AA.
AC Q9XDV0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit.
GN Name=pufL;
OS Erythrobacter sp. MBIC3960.
```



```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=94771;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MBIC3960;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hanada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR EMBL; AB027515; BAA78672.1; -; Genomic_DNA.
DR HSSP; P02954; 1YST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006119; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo_RC_1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
SQ SEQUENCE 278 AA; 30735 MW; 0B5618844B3C54FB CRC64;

Query Match 61.6%; Score 45; DB 2; Length 278;
Best Local Similarity 80.0%; Pred.No.26; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

QY 1 IEGPTLRQWL 10
DB 55 IEGPTLRPWL 64

RESULT 17
ID Q7WLX1_BORPA PRELIMINARY; PRT; 421 AA.
AC Q7WLX1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedLocustNames=BP0223;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30725.1; -; Genomic_DNA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45558 MW; A6CDBC98C731A49C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred.No.40; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 1;

QY 4 PTLROWLAAR 13
DB 221 PSLRDWAAR 230

RESULT 18
ID Q7WQUB_BORBR PRELIMINARY; PRT; 421 AA.
AC Q7WQUB;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative phenylacetate-CoA ligase.
GN OrderedLocustNames=BB0227;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30725.1; -; Genomic_DNA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR Complete proteome; Ligase.
SQ SEQUENCE 421 AA; 45558 MW; A6CDBC98C731A49C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 421;
Best Local Similarity 80.0%; Pred.No.40; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 1;

QY 4 PTLROWLAAR 13
DB 221 PSLRDWAAR 230

RESULT 19
ID Q885P2_PSESM PRELIMINARY; PRT; 756 AA.
AC Q885P2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Dimethylsulfoxide reductase.
GN OrderedLocustNames=PSPT01789;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.

```

```

OX NCBI_TaxID=323;
RN [1]
RC STRAIN=DC3000;
RA MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RX Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Uitterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowicz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL: AE016853; AAO55309.1; -; Genomic_DNA.
DR HSPF: Q57366; 1EUL.
DR TIGR: PSF01789; -.
DR GO: GO:0030151; F:molybdenum ion binding; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro: IPR006656; Molybdopterin.
DR InterPro: IPR006657; Mol. dinuc. bind.
DR Pfam: PF00384; Molybdopterin; 2.
DR Pfam: PF01568; Molybdopterin; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 756 AA; 83189 MW; 31E9614DEB22B2C CRC64;

Query Match 61.6%; Score 45; DB 2; Length 756;
Best Local Similarity 61.5%; Pred. No. 71;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAAAR 13
Db 222 LAGPTHQWLAVR 234
: ||| |||||
ID _GST3 CAEEL STANDARD; PRT; 207 AA.
AC 016156; Q21357;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 49, Last annotation update)
DE Glutathione S-transferase 3 (EC 2.5.1.18) (GST class-sigma) (CeGST3).
GN Name=Gst-3; ORFNames=K08F4.11;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC STRAIN=Bristol N2;
RA Tave W.N., Eschbach M.-L., Walter R.D., Henkle-Duehresen K.;
RT "Parasitoid mediates differential gene expression in C. elegans.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RC STRAIN=Bristol N2;
RA MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RC SEQUENCE REVISION.
RG WormBase consortium;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC exogenous and endogenous hydrophobic electrophiles (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
```

```

CC -1- SIMILARITY: Belongs to the GST superfamily. Sigma family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: AF010241; AAB65419.1; -; mRNA.
DR EMBL: Z68879; CAA93088.2; -; Genomic_DNA.
DR PIR: T37464; T37464.
DR HSPF: P24472; 1GUK.
DR EMBL: K08F4.11; Caenorhabditis elegans.
DR WormBase: WBGene00001751; gsc-3.
DR WormPept: K08F4.11; CE25050.
DR GO: GO:0004364; F:glutathione transferase activity; ISS.
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR004045; GST_Nterm.
DR InterPro: IPR012336; Thioredoxin-like.
DR InterPro: IPR012335; Thioredoxin_fold.
DR Pfam: PF00043; GST_C; 1.
DR Pfam: PF02798; GST_N; 1.
DR Complete proteome; Transferase.
KW Complete proteome; Transferase.
SQ SEQUENCE 207 AA; 23735 MW; 72545319FCFCEDBA CRC64;

Query Match 60.3%; Score 44; DB 1; Length 207;
Best Local Similarity 61.5%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAAAR 13
Db 190 IETPKLEWLAKR 202
: ||| |||||
ID _ECOS7 STANDARD; PRT; 536 AA.
AC Q8XBV3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Enterobactin synthetase component E (Enterobactin synthase E)
DE [includes: 2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)
DE (dihydroxybenzoic acid-activating enzyme); S-
DE dihydroxybenzoyltransferase (EC 2.3.1.-)];
GN Name=entB; OrderedLocusNames=z0736, EC0633;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
RA Roe D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink N.W., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-T., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobo T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
```

```

CC -!- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
CC benzoate (2,3-DHB) via ATP-dependent ppi exchange reactions, to
CC the acyladenylate. Then, catalyzes the acylation of holo-entB with
CC 2,3-DHB adenylate, preparing that molecule for amide bond
CC formation with L-serine (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
CC (2,3-dihydroxybenzoyl)adenylate.
CC -!- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)adenylate + holo-entB =
CC adenosine 5'-monophosphate + acyl-holo-entB.
CC -!- PATHWAY: Siderophore biosynthesis; enterobactin biosynthesis.
CC -!- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
CC complex called enterobactin synthetase (By similarity).
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family. EntE subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB005174; AAG54929.1; -; Genomic_DNA.
CC EMBL; BA000007; BAB34056.1; -; Genomic_DNA.
CC PIR; A99708; A99708.
CC HSSP; P40871; 1MD9.
CC InterPro; IPR000873; AMP-bind.
CC InterPro; IPR011963; DHB AMP_lig.
CC Pfam; PF00501; AMP-binding; 1.
CC TIGRFAMs; TIGR02275; DHB AMP_lig; 1.
CC PROSITE; PS00455; AMP BINDING; 1.
CC Acyltransferase; Complete proteome; Enterobactin biosynthesis; Iron;
CC Iron transport; Ligase; Multifunctional enzyme; Transferase;
CC Transport.
CC SEQUENCE 536 AA; 59040 MW; ABC8EB03209940A5 CRC64;
CC -----
Query Match 60.3%; Score 44; DB 1; Length 536;
Best Local Similarity 57.1%; Pred. No. 75;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 IEGPTLRQWLAAARA 14
Db 521 VDKKQLRWLASRA 534
:: |||||::|
-----
RESULT 22
ENTE ECOLI STANDARD; PRT; 536 AA.
AC P10378; P15049; P77773;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Enterobactin synthetase component E (Enterochelin synthase E)
DE [includes: 2,3-dihydroxybenzoate-AMP ligase (EC 2.7.7.58)
DE (Dihydroxybenzoic acid-activating enzyme); S-
DE dihydroxybenzoyltransferase (EC 2.3.1.-)].
GN Name=entE; Ordered locus names=b0594;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=89290355; PubMed=2525505; DOI=10.1016/0378-1097(89)90450-3;
RA Staab J.P., Elkins M.F., Earhart C.F.;
RT "Nucleotide sequence of the Escherichia coli entE gene.";
RL FEMS Microbiol. Lett. 50:15-19(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

```

```

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RA Chung E., Allen E., Araujo R., Aparicio A.M., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Iew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RT "Sequence of minutes 4-25 of Escherichia coli.";
RN Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RL [4]
RP NUCLEOTIDE SEQUENCE OF 1-8.
RC STRAIN=K12;
RX MEDLINE=90236256; PubMed=2110093; DOI=10.1016/0378-1097(90)90120-F;
RA Elkins M.F., Earhart C.F.;
RT "Opacity factor from group A streptococci is an apoproteinase.";
RL FEMS Microbiol. Lett. 56:35-40(1988).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 393-536.
RC MEDLINE=89123155; PubMed=2521622;
RA Liu J., Duncan K., Walsh C.T.;
RT "Nucleotide sequence of a cluster of Escherichia coli enterobactin
RT biosynthesis genes: identification of entA and purification of its
RT product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase.";
RL J. Bacteriol. 171:791-798(1989).
RN [6]
RP FUNCTION.
RX MEDLINE=97361959; PubMed=9214294; DOI=10.1021/bi970453p;
RA Gehring A.M., Bradley K.A., Walsh C.T.;
RT "Enterobactin biosynthesis in Escherichia coli: isochorismate lyase
RT (EntB) is a bifunctional enzyme that is phosphopantetheinylated by
RT EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate.";
RL Biochemistry 36:8495-8503(1997).
CC -!- FUNCTION: Activates the carboxylate group of 2,3-dihydroxy-
CC benzoate (2,3-DHB), via ATP-dependent ppi exchange reactions, to
CC the acyladenylate. Then, catalyzes the acylation of holo-entB with
CC 2,3-DHB adenylate, preparing that molecule for amide bond
CC formation with L-serine.
CC -!- CATALYTIC ACTIVITY: ATP + 2,3-dihydroxybenzoate = diphosphate +
CC (2,3-dihydroxybenzoyl)adenylate.
CC -!- CATALYTIC ACTIVITY: (2,3-dihydroxybenzoyl)adenylate + holo-entB =
CC adenosine 5'-monophosphate + acyl-holo-entB.
CC -!- PATHWAY: Siderophore biosynthesis; enterobactin biosynthesis.
CC -!- SUBUNIT: Proteins entB, entD, entE, and entF form a multienzyme
CC complex called enterobactin synthetase.
CC -!- INTERACTION:
CC P33371:dsuC; NbExp=1; IntAct=EBI-550322, EBI-562765;
CC P45523:fkpA; NbExp=1; IntAct=EBI-550322, EBI-550304;
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family. EntE subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U00096; AAC73695.1; -; Genomic DNA.
CC EMBL; U02598; AAB40794.1; -; Genomic DNA.
CC EMBL; X15058; CAA33158.1; -; Genomic DNA.
CC EMBL; M24148; AAL16101.1; -; Unassigned DNA.
CC EMBL; M36700; AAL18492.1; -; Genomic DNA.
CC PIR; H64792; SYCECB.
CC HSSP; P40871; 1MD9.
CC InterPro; IPR011963; DHB AMP_lig.
CC EcoGene; EG10263; entE.
CC InterPro; IPR000873; AMP-bind.
CC InterPro; IPR011963; DHB AMP_lig.
CC Pfam; PF00501; AMP-binding; 1.
CC PRINTS; PR00154; AMPBINDING.

```

DR TIGRFAMs: TIGR02275; DHB AMP lig; 1.
 DR PROSITE; P800455; AMP BINDING; 1.
 KW Acyltransferase; Complete proteome; Enterobactin biosynthesis; Iron;
 KW Iron transport; Ligase; Multifunctional enzyme; Transferase;
 FT TRANSPORT.
 FT CONFLICT 369 378 DAEGNPLPQG -> ECRKSTAAAR (in Ref. 1).
 SQ SEQUENCE 536 AA; 59112 MW; F818942DFDD8DC99 CRC64;

Query Match 60.3%; Score 44; DB 1; Length 536;
 Best Local Similarity 57.1%; Pred. No. 75;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 Db 521 VDKKQLRQWLASRA 534

RESULT 23

ID Q83M10_SHIFL PRELIMINARY; PRT; 536 AA.
 AC Q83M10; Q7C283;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE 2,3-dihydroxybenzoate-AMP ligase
 GN Name=entE; OrderedLocNames=S0514, SF0508;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OC NCBI_TaxID=623;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RX DOI=10.1128/TAI.71.5.2775-2786.2003;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -I- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.

DR EMBL; AE005674; AAN42156.1; -; Genomic DNA.
 DR EMBL; AE016979; AAP16028.1; -; Genomic DNA.
 DR HSSP; P40871; 1MDB.
 DR GO; GO:0016874; P:ligase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR011963; DHB AMP lig.
 DR Pfam; PF00501; AMP-binding; I.
 DR PRINTS; PR00154; AMPBINDING.
 DR TIGRFAMs; TIGR02275; DHB AMP lig; 1.
 DR PROSITE; P800455; AMP BINDING; UNKNOWN_1.
 KW Complete proteome; Ligase.
 SQ SEQUENCE 536 AA; 58851 MW; ABABD6BB8692A8D2 CRC64;

Query Match 60.3%; Score 44; DB 2; Length 536;
 Best Local Similarity 57.1%; Pred. No. 75;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAARA 14
 Db 521 VDKKQLRQWLASRA 534

RESULT 24

ID Q4ZQD5_PSESY PRELIMINARY; PRT; 756 AA.
 AC Q4ZQD5;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Trimethylamine-N-oxide reductase (Cytochrome c) (EC 1.7.2.3).
 GN ORFNames=Psyr_3605;
 OS Pseudomonas syringae pv. syringae B728a.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=205918;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=B728a;
 RG DOE Joint Genome Institute;
 RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
 RA Nolan M., Goltsman E., Thiel J., Malfatti S., Lapidus A., Dettler J.C.,
 RA Land M., Richardson P.M., Kyrpides N.C., Ivanova N.;
 RT "Comparison of two complete genome sequences of Pseudomonas syringae
 RT pv. syringae B728a and pv. tomato DC3000";
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=B728a;
 RA Loper J.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=B728a;
 RA Feil H., Feil W.S., Lindow S.E.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CP000075; AAY38637.1; -; Genomic DNA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0050626; F:trimethylamine-N-oxide reductase (cytochrom. .); IEA.
 KW Oxidoreductase.

SQ SEQUENCE 756 AA; 83765 MW; 892D050625BA61DE CRC64;
 Query Match 60.3%; Score 44; DB 2; Length 756;
 Best Local Similarity 72.7%; Pred. No. 1.1e+02;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAR 13
 Db 224 GPTHQWLAVR 234

RESULT 25

ID Q7G4N4_ORYSA PRELIMINARY; PRT; 760 AA.
 AC Q7G4N4;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Putative gag-pol polypeptide (Hypothetical protein).
 GN ORFNames=OSUNB003BH12.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";

```
RL Science 300:1566-1569 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RN Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,
RA Bera J., Kim M., Jin S., Fadrosch D., Vuong H., Overton II L.,
RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,
RA de Vazeilles A., White O., Salzberg S., Fraser C.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017067; AAP52546.1; -; Genomic_DNA.
DR EMBL; AC105932; AAX95563.1; -; Genomic_DNA.
DR Gramene; Q8LMK9; -.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR Hypothetical protein; Polypotein.
RW KQ
SQ SEQUENCE 760 AA; 82020 MW; C51F91AA2EB32A28 CRC64;
Query Match 60.3%; Score 44; DB 2; Length 760;
Best Local Similarity 46.2%; Pred. No. 1.1e-02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 13
DB 661 LHGPTLQHWMAVK 673

RESULT 26
KSGA_TREPA
ID KSGA_TREPA STANDARD; PRT; 285 AA.
AC O83357;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Dimethyladenosine transferase (EC 2.1.1.-) (S-adenosylmethionine-6-N',
DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High
DE level kasugamycin resistance protein ksgA) (kasugamycin
DE dimethyltransferase).
DE Name=ksgA; OrderedLocusNames=TP0337;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwin M.B., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artlich P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388 (1998).
CC -1- FUNCTION: Specifically dimethylates two adjacent adenosines in the
CC loop of a conserved hairpin near the 3' end of 16S rRNA in the 30S
CC particle. Its inactivation leads to kasugamycin resistance (By
CC similarity).
CC -1- SIMILARITY: Belongs to the rRNA adenine N-6-methyltransferase
CC family. KsgA subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE001213; AAC65323.1; -; Genomic_DNA.
DR PIR; G71337; G71337.
DR TIGR; TP0337; -.
DR HAMAP; MF 00607; -; 1.
DR InterPro; IPR011530; ksgA.
DR InterPro; IPR001737; rRNA_meth_trans.
DR InterPro; IPR000051; SAM_Ld.
DR PANTHER; PTHR11727:SF6; ksgA; 1.
DR PANTHER; PTHR11727:RMA_meth_trans; 1.
DR Pfam; PF00398; RnaAD; 1.
DR SMART; SM00650; RAD; 1.
DR TIGRFAMs; TIGR00755; ksgA; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
RW Antibiotic resistance; Complete proteome; Methyltransferase;
KW rRNA processing; Transferase.
SQ SEQUENCE 285 AA; 32276 MW; 3AF08CBEL16B5D4F CRC64;
Query Match 58.9%; Score 43; DB 1; Length 285;
Best Local Similarity 64.3%; Pred. No. 59;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEGPTLRQWLAAR 14
DB 98 IEGDVLQQWHA AAA 111

RESULT 27
Q7UQE4_RHOBA
ID Q7UQE4_RHOBA PRELIMINARY; PRT; 297 AA.
AC Q7UQE4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
DE OrderedLocusNames=RB6375;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Firellula.
OX NCBI_TaxID=1117;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=2735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Firellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294144; CAD74759.1; -; Genomic_DNA.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR003169; GYF.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50829; GYF; 1.
RW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475F670F02C78E9B CRC64;
Query Match 58.9%; Score 43; DB 2; Length 297;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 EGPTRLQWMA 11
DB 176 DGPTRKQWIS 185

RESULT 28
Q8EJ00_SHEON
ID Q8EJ00_SHEON PRELIMINARY; PRT; 306 AA.
```

```

AC Q8EJ00;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Prophase Mus01, major head subunit, putative.
GN OrderedLocusNames=SQ0675;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Unayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AB015513; AAN53753.1; -; Genomic_DNA.
DR TIGR; SO0675; -.
KW Complete proteome.
SQ SEQUENCE 306 AA; 34370 MW; F54CCA118AA288CB CRC64;

Query Match 58.9%; Score 43; DB 2; Length 306;
Best Local Similarity 60.0%; Pred. No. 64;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTLRQWLAAR 13
Db 54 PTMRWIGAR 63
||||:|||||

RESULT 29
Q8ZVT5 PYRAE
ID Q8ZVT5 PYRAE PRELIMINARY; PRT; 354 AA.
AC Q8ZVT5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0634.
GN OrderedLocusNames=PAE0634;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AB009776; AAL62908.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 354 AA; 38642 MW; C5799F975B972941 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 354;
Best Local Similarity 61.5%; Pred. No. 74;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWLAAR 13
Db 84 IDRFQLEQWLAAR 96
||||:|||||

```

```

RESULT 30
Q82PX5 STRAW
ID Q82PX5 STRAW PRELIMINARY; PRT; 377 AA.
AC Q82PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV747;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
DR EMBL; BA000030; BAC68457.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 377 AA; 41307 MW; 0253176AAAB62F3 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 377;
Best Local Similarity 61.5%; Pred. No. 79;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWLAAR 13
Db 168 MEGPDRLRAWLPNR 180
||||:|||||

RESULT 31
Q83436 TREPA
ID Q83436 TREPA PRELIMINARY; PRT; 683 AA.
AC Q83436;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein TP0421.
GN OrderedLocusNames=TP0421;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khakhal H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."

```

```

RL Science 281:375-388(1998).
DR EMBL; AE001220; AAC65409.1; -; Genomic_DNA.
DR PIR; B71325; B71325.
DR TIGR; TP0421; -.
DR GO; GO:0005488; F:binding; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR013017; NHL_rep.
DR InterPro; IPR001440; TPR_rep.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF01436; NHL; 5.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome; Repeat; TPR repeat.
SQ SEQUENCE 683 AA; 74519 MW; F91407FA7094AAD1 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 683;
Best Local Similarity 69.2%; Pred. No. 1.4e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAAAR 13
Db 89 IEGALHQWGAAR 101
|||||
89 IEGALHQWGAAR 101

RESULT 32
Q617Q3 CAEBR
ID Q617Q3 CAEBR PRELIMINARY; PRT; 728 AA.
AC Q617Q3
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14874.
GN Name=CBG14874;
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=62338;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; CAAC0100068; CAE68902.1; -; Genomic_DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Complete proteome; Hypothetical protein; Metal-binding;
Ubl conjugation pathway; Zinc; Zinc-finger.
SQ SEQUENCE 754 AA; 85324 MW; 41BAA9297FA3BF05 CRC64;

Query Match 58.9%; Score 43; DB 2; Length 754;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAAAR 13
Db 626 GPCLRWLAVK 636
|||||
626 GPCLRWLAVK 636

RESULT 34
Q527F0 MAGGR
ID Q527F0 MAGGR PRELIMINARY; PRT; 1157 AA.
AC Q527F0
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG05740.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gvaltzen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,

```


Query Match 57.5%; Score 42; DB 2; Length 126;
Best Local Similarity 58.3%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GPTLRQWLAA 14
Db 108 GPDLRWAGSRA 119

RESULT 37
Q4SMZ4 TETNG
ID Q4SMZ4 TETNG PRELIMINARY; PRT; 127 AA.
AC Q4SMZ4;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Chromosome 6 SCAP14544, whole genome shotgun sequence.
GN ORFNames=GSTNG00015515001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn W., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAG01014544; CAF97988.1; -: Genomic DNA.
SQ SEQUENCE 127 AA; 14785 MW; 0DE1B75D2B88BC8 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 127;
Best Local Similarity 57.1%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAA 14
Db 59 IEDPALTRWVHARA 72

RESULT 38
Q8XPQ9 RALSO
ID Q8XPQ9 RALSO PRELIMINARY; PRT; 252 AA.
AC Q8XPQ9;
DT 01-WAR-2002 (TRENBLrel. 20, Created)
DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-WAR-2004 (TRENBLrel. 26, Last annotation update)
DE PUTATIVE TRANSCRIPTION REGULATOR PROTEIN.
GN OrderedLocusNames=RS02135;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.

NCBI_TaxID=305;
[1]
NUCLEOTIDE SEQUENCE.
STRAIN=GMI1000;
MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).
CC -!- SIMILARITY: Contains 1 HTH luxR-type DNA-binding domain.
DR EMBL; AL646085; CAD18730.1; -: Genomic_DNA.
DR HSSP; P11470; 1FSE.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00196; GcrE; 1.
DR PRINTS; PR01590; HTHFIS.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00421; HTH_LuxR; 1.
KW Complete proteome; DNA-binding; Plasmid; Transcription;
KW Transcription regulation.
SQ SEQUENCE 252 AA; 27666 MW; 483403EE3267C2E CRC64;

Query Match 57.5%; Score 42; DB 2; Length 252;
Best Local Similarity 53.8%; Pred. No. 78;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IEGPTLRQWLAA 13
Db 76 IDTPMRWLATR 88

RESULT 39
P90433 SIVCZ
ID P90433 SIVCZ PRELIMINARY; PRT; 313 AA.
AC P90433;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Truncated reverse transcriptase (fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Smith J.M., Kraiselburd E.N., Torres J.V.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U83413; AAB41428.1; -: Genomic_DNA.
DR HSSP; Q07387; 1TCW.
DR SMR; P90433; 69-167, 173-313.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartic_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.

```

DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
DR KW Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase.
FT NON_TER 1
SQ SEQUENCE 313 AA; 34675 MW; 5A0BB016783FC8A6 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 313;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

RESULT 40
Q85SN9_9CAUD PRELIMINARY; PRT; 325 AA.
AC Q85SN9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp74.
OS Mycobacteriophage Che9d.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205876;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2259660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AY129336; AAN07992.1; -; Genomic DNA.
SQ SEQUENCE 325 AA; 35999 MW; 04265796D0B4FC1D CRC64;

Query Match 57.5%; Score 42; DB 2; Length 325;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAARA 14
Db 302 DGPTVQEAARA 314

RESULT 41
P95613_RHIGA
ID P95613_RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium gallegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAMBI.
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstroem K.;
RL Submitted (Oct-1996) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: Contains 1 HTH lyase-type DNA-binding domain.
DR EMBL; Y08963; CAA70157.1; -; Genomic DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

```

```

DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_Lyase.
DR InterPro; IPR005119; LyseR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LyseR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00931; HTH_LYSE; 1.
KW Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 57.5%; Score 42; DB 2; Length 326;
Best Local Similarity 50.8%; Pred. No. 1e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQWLAAR 13
Db 204 KGPSLEQWLSSQ 215

RESULT 42
Q7FAN4_ORYSA
ID Q7FAN4_ORYSA PRELIMINARY; PRT; 375 AA.
AC Q7FAN4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE OSUNB0060E08.2 protein.
GN Name=OSUNB0060E08.2;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=1247439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Xu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu Z., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606669; CAE04739.1; -; Genomic DNA.
DR Gramene; Q7XPP6; -.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR00210; BTB_POZ.
DR InterPro; IPR02083; MATH.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00917; MATH; 1.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00661; MATH; 1.
DR PROSITE; PS00097; BTB; 1.
DR PROSITE; PS0144; MATH; 1.
DR PROSITE; PS0144; MATH; 1.
SQ SEQUENCE 375 AA; 41043 MW; 20FC6E99E4750816 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEPTLRQWLAAR 13
Db 138 MERFLRQWLLR 150

RESULT 43
Q7XPP6_ORYSA

```

```

ID Q7XPP6 ORYSA PRELIMINARY; PRT; 375 AA.
AC Q7XPP6
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNBa0053K19.27 protein.
GN Name=OSJNBa0053K19.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606645; CAE03519.2; -; Genomic_DNA.
DR Gene; Q7XPP6; -.
DR GO; GO:000515; F:protein binding; IEA.
DR InterPro; IPR000210; BTP_POZ.
DR InterPro; IPR002083; MATH.
DR Pfam; PF00651; BTP; 1.
DR SMART; SM00917; MATH; 1.
DR SMART; SM00225; BTP; 1.
DR SMART; SM00061; MATH; 1.
DR PROSITE; PS50097; BTP; 1.
DR PROSITE; PS50144; MATH; 1.
SQ SEQUENCE 375 AA; 41043 MW; 20FC6E99E4750816 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 375;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IEQPTLRQWL 13
Db 138 MERPLRLQWLRR 150

RESULT 44
Q9SLB9 ARATH PRELIMINARY; PRT; 450 AA.
AC Q9SLB9; Q9ASU2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Expressed protein (At2g42400/MK10.12) (transcription factor
DE AtVOZ2).
GN Name=At2g42400; Synonyms=AtVOZ2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

Town C.D., Kaul S.;
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[3]
NUCLEOTIDE SEQUENCE.
Chen R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
Palm C.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[4]
NUCLEOTIDE SEQUENCE.
Chen R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
Theologis A., Ecker J.R.;
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[5]
NUCLEOTIDE SEQUENCE.
Mitsuda N., Hisabori T., Takeyasu K., Sato M.H.;
Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005956; AAD23723.2; -; Genomic_DNA.
DR EMBL; AY078048; AAL77749.1; -; mRNA.
DR EMBL; AF361831; AAK32843.1; -; mRNA.
DR EMBL; AB125257; BADI7858.1; -; mRNA.
DR PIR; E84853; E84853.
SQ SEQUENCE 450 AA; 50566 MW; 44DBE7B4AB689B95 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 450;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 IEQPTLRQWL 10
Db 320 VEGETIREWL 329

RESULT 45
Q5GW75 XANOR PRELIMINARY; PRT; 532 AA.
AC Q5GW75;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Phosphoanhydride phosphohydrolase.
GN Name=appa; OrderedLocNames=XOO3792;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KXO85;
RX PubMed=156737118; DOI=10.1093/nar/gki206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
the bacterial blight pathogen of rice."
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AE013598; AAW77046.1; -; Genomic_DNA.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
Complete proteome; Hydrolase.
```

```

SQ SEQUENCE 532 AA; 55188 MW; 56D054783C4B3392 CRC64;
Query Match 57.5%; Score 42; DB 2; Length 532;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWLAARA 14
: |||||:
Db 203 LSGBYLRQWLAQA 216

RESULT 46
Q9N6P9 LEIMA
ID Q9N6P9 LEIMA PRELIMINARY; PRT; 586 AA.
AC Q9N6P9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE L354.2 (Hypothetical protein)
GN Name=L354.2; ORFNames=LMJ_1272;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Myler P.J., Sisk E., Cawthra J., Handley F., Vogt C., Robertson L.,
RA McDonagh P., Ivans A., Nguyen D., Munden H., Stuart K.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RX PubMed=16020728; DOI=10.1126/science.1112680;
RA Ivans A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA Bertrinan M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA Apostolou Z., Attipoe P., Bacon N., Bauser C., Beck A., Beverley S.M.,
RA Blanchettin G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
RA Ceadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
RA Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA Fazellina G., Foaker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
RA Goble A., Goffeau A., Harris D., Hertz-Powler C., Hilbert H., Horn D.,
RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
RA Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA Oliver K., O'Neill S., Pencony M., Pohl T.M., Price C., Furnelle B.,
RA Robil M.A., Rabinowitch E., Reinhardt R., Rieger M., Rinta J.,
RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
RA Myler P.J.;
RT "The Genome of the Kinetoplastid Parasite, Leishmania major.";
RL Science 309:436-442(2005)
DR EMBL; AC022473; AAF69556.1; -; Genomic DNA.
DR EMBL; CP000081; AA214566.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 586 AA; 63995 MW; 208367C213896AF3 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 586;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IEQPTLRQWLAARA 14
: |||||:
Db 58 VEAPLITQWMTAAA 71

RESULT 47
POL_SIVS4
ID POL_SIVS4 STANDARD; PRT; 1019 AA.
AC P12502;
DT 01-OCT-1989 (Rel. 12, Created)

```

```

DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.-);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)].
GN Name=POL;
OS Simian immunodeficiency virus (isolate F236/smH4) (SIV-SM).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11737;
[1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89262053; PubMed=2786147; DOI=10.1038/339389a0;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVam) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -1- FUNCTION: During replicative cycle of retroviruses, the reverse-
CC transcribed viral DNA is integrated into the host chromosome by
CC the viral integrase enzyme. RNase H activity is associated with
CC the reverse transcriptase.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoster.
CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
CC diphosphate + DNA(n+1).
CC -1- PFM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC -1- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC -1- SIMILARITY: Contains 1 integrase catalytic domain.
CC -1- SIMILARITY: Contains 1 integrase-type DNA-binding domain.
CC -1- SIMILARITY: Contains 1 integrase-type zinc finger.
CC -1- SIMILARITY: Contains 1 peptidase A2 domain.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC -1- SIMILARITY: Contains 1 RNase H domain.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; X14307; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR HSP; P04584; 1MU2.
DR SWP; P12502; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR HIV; X14307; POL$SMWH4.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR AIDS; Aspartyl protease; DNA integration; DNA recombination;

```

KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
 KW Nuclease; Nucleosidyltransferase; Polyprotein; Protease;
 KW RNA-directed DNA polymerase; Transferase; Zinc; Zinc-finger.
 FT CHAIN 1 167
 FT DOMAIN 88 157 Peptidase A2.
 FT DOMAIN 211 401 Reverse transcriptase.
 FT DOMAIN 600 723 RNase H.
 FT DOMAIN 779 930 Integrase catalytic.
 FT ZN_FING 729 770 Integrase-type.
 FT ACT BIND 949 996 Integrase-type.
 FT ACT SITE 93 By similarity.
 SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match 57.5%; Score 42; DB 1; Length 1019;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 EGPTLRQW 9
 ||| ||||
 DB 184 EGPKLRQW 191

RESULT 48

ID P89154_SIVCZ PRELIMINARY; PRT; 1019 AA.
 AC P89154;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Pol protein (Fragment).
 GN Name=pol;
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Lentivirus; Primate lentivirus group.
 OX NCBI_TaxID=11723;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SiVemE543;
 RC MEDLINE=97115152; PubMed=8995688;
 RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,
 RA Elkins W.R., Montefiori D.C.;
 RT "A molecularly cloned, pathogenic, neutralization-resistant simian
 immunodeficiency virus, SiVemE543-3.";
 RL J. Virol. 71:1608-1620(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SiVemE543;
 RA Ourmanov I.K., Dehghani H., Kuwata T., Hirsch V.M.;
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
 the viral RNA genome into double-stranded viral DNA. The RNase H
 domain of the reverse transcriptase performs two functions. It
 degrades the RNA template and specifically removes the RNA primer
 from the RNA/DNA hybrid. Following nuclear import, the integrase
 catalyzes the insertion of the linear, double-stranded viral DNA
 into the host cell chromosome. Endogenous Pol proteins may have
 kept, lost or modified their original function during evolution
 (By similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
 conserved among many retroviruses (By similarity).
 CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; U72748; AAC56559.2; -; Genomic_DNA.
 DR SMR; P89154; 69-167, 781-939.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_N_Zn_bd.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR InterPro; IPR009007; Pept_Aspartc_cat.
 DR InterPro; IPR001995; Peptidase_A2_cat.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR012337; RNaseH_fold.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR010659; RVT_Connect.

DR InterPro; IPR010661; RVT_thumb.
 DR InterPro; IPR000477; RVTse.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF00552; Integrase_1.
 DR Pfam; PF02022; Integrase_Zn; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00665; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT_1; 1.
 DR Pfam; PF06815; RVT_connect; 1.
 DR Pfam; PF06817; RVT_thumb; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS00994; INTEGRASE; 1.
 DR PROSITE; PS1027; INTEGRASE_DBD; 1.
 DR PROSITE; PS0879; RNase H; 1.
 DR PROSITE; PS0878; RT_POL; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
 DR PROSITE; PS0876; ZF-INTEGRASE; 1.
 KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
 KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 EGPTLRQW 9
 ||| ||||
 DB 184 EGPKLRQW 191

RESULT 49

ID Q7ZBR7_SIVCZ PRELIMINARY; PRT; 1019 AA.
 AC Q7ZBR7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Pol (Fragment).
 GN Name=pol;
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Lentivirus; Primate lentivirus group.
 OX NCBI_TaxID=11723;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=22628501; PubMed=12743298;
 RX DOI=10.1128/JVI.77.11.6405-6418.2003;
 RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
 RT "Unique pattern of convergent envelope evolution in simian
 immunodeficiency virus-infected rapid progressor macaques: association
 with CD4-independent usage of CCR5.";
 RL J. Virol. 77:6405-6418(2003).
 CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
 the viral RNA genome into double-stranded viral DNA. The RNase H
 domain of the reverse transcriptase performs two functions. It
 degrades the RNA template and specifically removes the RNA primer
 from the RNA/DNA hybrid. Following nuclear import, the integrase
 catalyzes the insertion of the linear, double-stranded viral DNA
 into the host cell chromosome. Endogenous Pol proteins may have
 kept, lost or modified their original function during evolution
 (By similarity).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
 conserved among many retroviruses (By similarity).
 CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; AY221514; AAO67307.1; -; Genomic_DNA.
 DR HSP; P04584; 1M02.
 DR SMR; Q7ZBR7; 69-167, 781-939.
 DR MEROPS; A02.002; -.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.

```

DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:Integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA transposition; IEA.
DR GO; GO:0006313; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase_1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115341 MW; A886525DF1BE26F CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

```

```

RT GO; GO:0003677; F:DNA binding; IEA.
RT GO; GO:0008907; F:Integrase activity; IEA.
RL J. Virol. 77:6405-6418(2003).
CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The LPOG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (by similarity).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221515; AA067309.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; O7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:Integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA transposition; IEA.
DR GO; GO:0006313; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH_1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EGPTLRQW 9
Db 184 EGPKLRQW 191

```

"Unique pattern of convergent envelope evolution in simian

Search completed: May 12, 2006, 10:52:00
Job time : 114.077 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 65.3846 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	92.9	11	2	Aay22365 TPO recep
2	26	92.9	11	2	Aay22370 TPO recep
3	26	92.9	11	3	Aab16986 TPO-mimet
4	26	92.9	11	3	Aab16988 TPO-mimet
5	26	92.9	11	5	Abb72872 TPO mimet
6	26	92.9	11	5	Abb72874 TPO mimet
7	26	92.9	11	7	Adj73023 TPO mimet
8	26	92.9	11	7	Adj73025 TPO mimet
9	26	92.9	11	8	Adj52658 CH1 delet
10	26	92.9	11	8	Adj52660 CH1 delet
11	26	92.9	11	8	Adj51619 CH1 delet
12	26	92.9	11	8	Adj51621 CH1 delet
13	26	92.9	13	2	Aay22407 TPO recep
14	26	92.9	95	3	Aay65258 Human 5'
15	26	92.9	95	8	Adu72822 Signal pe
16	26	92.9	95	9	Adz73813 Human inc
17	26	92.9	357	6	Abu48350 Protein e
18	26	92.9	465	8	Adi67225 Lactobaci
19	26	92.9	486	4	Abb71566 Drosophi
20	26	92.9	510	7	Abu66441 Klebsiell
21	26	92.9	2559	3	Abm96913 M. xanthu
22	25	89.3	9	3	Abb16977 TPO-mimet
23	25	89.3	9	5	Abb72863 TPO mimet
24	25	89.3	9	6	Abg71745 TPO recep

98	24	85.7	264	7	ADM04596	Adm04596 Human pro	171	24	85.7	655	8	ADJ49379	Oil-assoc
99	24	85.7	266	3	RAG34624	Rag34624 Arabidops	172	24	85.7	664	7	ADJ69530	Human hea
100	24	85.7	267	7	ABO63689	AbO63689 Klebsiell	173	24	85.7	666	7	ABO83796	Pseudomon
101	24	85.7	292	4	AAB80103	Aab80103 Corynebac	174	24	85.7	698	8	ADI19358	Adi19358 soybean t
102	24	85.7	304	6	ABU16626	Abu16626 Protein e	175	24	85.7	698	8	ADJ65943	Abj65943 soybean t
103	24	85.7	315	3	RAG26425	Rag26425 Arabidops	176	24	85.7	710	6	ABR58651	AbR58651 Human can
104	24	85.7	315	3	RAG46503	Rag46503 Arabidops	177	24	85.7	723	6	ABU24903	Protein e
105	24	85.7	316	5	ADI18829	Adi18829 Murine NO	178	24	85.7	737	8	ADR10261	ADR10261 Human pro
106	24	85.7	316	5	ABW02143	AbW02143 Human GPC	179	24	85.7	741	6	ABU21306	Protein e
107	24	85.7	317	2	AAR45363	Aar45363 Naei rest	180	24	85.7	802	6	ADA54442	Human pro
108	24	85.7	324	5	ABP56427	AbP56427 Human glu	181	24	85.7	874	4	ABG26108	Novel hum
109	24	85.7	342	9	ABE41377	AbE41377 L. pneumo	182	24	85.7	891	3	AAB26596	Synechocy
110	24	85.7	345	3	RAG26424	Rag26424 Arabidops	183	24	85.7	897	6	ADB11560	Allioiococ
111	24	85.7	345	3	RAG46502	Rag46502 Arabidops	184	24	85.7	924	8	ADP99020	C. albica
112	24	85.7	350	4	AAB50667	Aab50667 C. elegan	185	24	85.7	966	6	ABM70145	Photorhab
113	24	85.7	351	4	AAB50666	Aab50666 C. elegan	186	24	85.7	993	4	ABB65429	Drosophil
114	24	85.7	351	9	ABR38085	AbR38085 L. pneumo	187	24	85.7	996	9	ABE21645	Adhesin-1
115	24	85.7	360	4	RAG91578	Rag91578 C glutami	188	24	85.7	997	4	ABG15941	Novel hum
116	24	85.7	363	3	RAG26423	Rag26423 Arabidops	189	24	85.7	1001	8	ADI28617	Mycoplasma
117	24	85.7	363	3	RAG46501	Rag46501 Arabidops	190	24	85.7	1166	8	ADN20571	Bacterial
118	24	85.7	363	5	ABR93830	AbR93830 Herbicida	191	24	85.7	2951	6	ABM67717	Photorhab
119	24	85.7	363	8	ADJ49252	AdJ49252 Oil-assoc	192	24	85.7	2975	6	ABM68759	Photorhab
120	24	85.7	363	8	ADJ49645	AdJ49645 Oil-assoc	193	24	85.7	3041	6	ABM70325	Photorhab
121	24	85.7	363	8	ADJ49605	AdJ49605 Oil-assoc	194	24	85.7	4560	6	ABM67454	Photorhab
122	24	85.7	385	3	RAG52926	Rag52926 Arabidops	195	23	82.1	10	2	AAJ22355	TPO recep
123	24	85.7	385	3	RAG41437	Rag41437 Arabidops	196	23	82.1	10	3	ABJ16979	TPO-mimet
124	24	85.7	430	6	ADB08462	AdB08462 Allioiococ	197	23	82.1	10	5	ABB72865	TPO mimet
125	24	85.7	430	8	ADJ27093	AdJ27093 Allioiococ	198	23	82.1	10	7	ADJ73016	TPO mimet
126	24	85.7	432	4	AAU51431	Aau51431 Propionib	199	23	82.1	10	8	ADJ52651	CH1 delet
127	24	85.7	432	6	ABM47950	AbM47950 Propionib	200	23	82.1	10	8	ADJ51612	CH1 delet
128	24	85.7	435	3	ABJ10917	AbJ10917 Human reg	201	23	82.1	12	5	ABB05304	Polyureth
129	24	85.7	441	7	ADM05130	AdM05130 Human pro	202	23	82.1	15	2	AAJ22395	TPO recep
130	24	85.7	456	6	ADB11562	AdB11562 Allioiococ	203	23	82.1	15	2	AAJ22403	TPO recep
131	24	85.7	461	3	RAG11836	Rag11836 Arabidops	204	23	82.1	15	2	AAJ22373	TPO recep
132	24	85.7	461	3	RAG46490	Rag46490 Arabidops	205	23	82.1	15	2	AAJ22390	TPO recep
133	24	85.7	462	3	RAG46489	Rag46489 Arabidops	206	23	82.1	15	2	AAJ22378	TPO recep
134	24	85.7	462	3	RAG11835	Rag11835 Arabidops	207	23	82.1	15	2	AAJ22397	TPO recep
135	24	85.7	464	8	ADX94959	AdX94959 Plant ful	208	23	82.1	25	4	ABM16749	Peptide #
136	24	85.7	465	3	RAG11834	Rag11834 Arabidops	209	23	82.1	25	4	ABB35735	Peptide #
137	24	85.7	465	3	RAG11834	Rag11834 Arabidops	210	23	82.1	25	4	AAJ29237	Peptide #
138	24	85.7	465	8	ADT55526	AdT55526 Plant pol	211	23	82.1	25	4	ABB30570	Peptide #
139	24	85.7	468	6	ABJ25769	AbJ25769 Aspergill	212	23	82.1	25	4	ABB21161	Peptide #
140	24	85.7	469	6	ABJ26369	AbJ26369 Aspergill	213	23	82.1	25	4	ABM68926	Human bon
141	24	85.7	471	3	AAE21094	AAe21094 Yeast HEL	214	23	82.1	25	4	ABM56544	Human bra
142	24	85.7	471	6	ABR52624	AbR52624 Protein s	215	23	82.1	25	4	ABG50591	Human liv
143	24	85.7	471	7	ADB84329	AdB84329 Yeast rep	216	23	82.1	25	5	ABG38507	Peptide #
144	24	85.7	471	7	ADK62264	AdK62264 Disease t	217	23	82.1	62	5	ABP32688	Human ORF
145	24	85.7	481	4	ABE71798	AbE71798 Drosophil	218	23	82.1	62	5	ADG79445	Human sec
146	24	85.7	481	4	AAE03541	AAe03541 Drosophil	219	23	82.1	62	5	ADG79538	Human sec
147	24	85.7	481	7	ADB84328	AdB84328 Drosophil	220	23	82.1	68	8	ABO57882	Human gen
148	24	85.7	498	5	ABP73641	AbP73641 Candida a	221	23	82.1	76	4	AAU66802	Propionib
149	24	85.7	505	6	ABR53169	AbR53169 Protein s	222	23	82.1	76	6	ABM63321	Propionib
150	24	85.7	505	7	ADK63188	AdK63188 Disease t	223	23	82.1	79	9	ABM92162	M. xanthu
151	24	85.7	520	7	ABO80725	AbO80725 Pseudomon	224	23	82.1	109	3	AAJ01169	Human sec
152	24	85.7	527	8	ADS23096	AdS23096 Bacterial	225	23	82.1	111	8	ADI57234	FEN1 domi
153	24	85.7	530	5	ABP69308	AbP69308 Human pol	226	23	82.1	114	3	ABM40430	Human ORF
154	24	85.7	543	8	ADV89911	AdV89911 Streptoco	227	23	82.1	114	5	ABP09147	Human ORF
155	24	85.7	543	8	ADV83301	AdV83301 Streptoco	228	23	82.1	118	4	ABG11928	Novel hum
156	24	85.7	543	8	ADV81164	AdV81164 Streptoco	229	23	82.1	136	3	AAJ96838	PCPB-DQVD
157	24	85.7	543	8	ADK67350	AdK67350 Plant ful	230	23	82.1	142	6	ABO08396	Allioiococ
158	24	85.7	551	2	AAW18790	AAw18790 Corrected	231	23	82.1	147	7	ABO66415	Klebsiell
159	24	85.7	572	8	ADX74519	AdX74519 Plant ful	232	23	82.1	153	4	AAO05794	Novel hum
160	24	85.7	585	4	ABB60953	AbB60953 Drosophil	233	23	82.1	154	7	ADA07510	Human sec
161	24	85.7	591	4	ABB63312	AbB63312 Drosophil	234	23	82.1	154	8	ADN41234	Novel hum
162	24	85.7	600	8	ADS29505	AdS29505 Bacterial	235	23	82.1	164	3	ABM44991	Human sec
163	24	85.7	600	9	ABM96218	AbM96218 M. xanthu	236	23	82.1	164	3	ABM44992	Human sec
164	24	85.7	624	5	ABP65819	AbP65819 Bifidobac	237	23	82.1	171	6	ABU26770	Protein e
165	24	85.7	628	8	ADN20030	AdN20030 Bacterial	238	23	82.1	172	2	AAJ35324	Chlamydia
166	24	85.7	645	7	ADC10136	AdC10136 Human NOV	239	23	82.1	177	6	ABU27123	Protein e
167	24	85.7	654	6	ABR53790	AbR53790 Protein s	240	23	82.1	182	6	ADA36082	Acinetoba
168	24	85.7	654	7	ADK62018	AdK62018 Disease t	241	23	82.1	192	8	ADX75919	Plant ful
169	24	85.7	654	8	ADJ50308	AdJ50308 Oil-assoc	242	23	82.1	206	3	AAJ07266	Arabidops
170	24	85.7	654	8	ADS43755	AdS43755 Bacterial	243	23	82.1				

244	23	82.1	212	9	ABE40014	Aeb40014	L. pneumo	317	23	82.1	403	8	ADQ98864	Adq98864	Antagonis
245	23	82.1	214	7	ADQ98874	Adc96874	E. faeciu	318	23	82.1	403	8	ADSL10466	Adsl10466	Human the
246	23	82.1	215	9	ABE36622	Aeb36622	L. pneumo	319	23	82.1	403	8	ADSL14402	Adsl14402	Mouse squ
247	23	82.1	218	6	ABU35228	Abu35228	Protein e	320	23	82.1	405	9	ABM95316	Abm95316	M. xanthu
248	23	82.1	220	9	ABE37357	Aeb37357	L. pneumo	321	23	82.1	408	4	AAB48304	Abm48304	Human 2F2
249	23	82.1	220	9	ABE40677	Aeb40677	L. pneumo	322	23	82.1	409	2	AAR25693	Aar25693	Heat resi
250	23	82.1	220	9	ABE40670	Aeb40670	L. pneumo	323	23	82.1	409	2	AAW03544	Aaw03544	Bacillus
251	23	82.1	224	9	ABE37350	Aeb37350	L. pneumo	324	23	82.1	410	8	ADR10146	Adr10146	Human pro
252	23	82.1	226	4	ABG02211	Abg02211	Novel hum	325	23	82.1	411	4	ABG13838	Abg13838	Novel hum
253	23	82.1	237	3	ABG36022	Abg36022	Zea mays	326	23	82.1	414	5	ABG70045	Abg70045	Human pre
254	23	82.1	245	3	ABG07265	Aag07265	Arabidops	327	23	82.1	417	7	ABO73714	AbO73714	Pseudomon
255	23	82.1	246	2	ABY37499	Aay37499	Protein w	328	23	82.1	418	8	ABG13105	Abg13105	Bacterial
256	23	82.1	247	3	ABE42519	Aab42519	Human ORF	329	23	82.1	419	4	AAAM39925	Aam39925	Human pol
257	23	82.1	254	3	ABG20958	Aag20958	Arabidops	330	23	82.1	420	4	ADS24829	Ads24829	Bacterial
258	23	82.1	265	5	ABJ04666	Abj04666	Mycobacte	331	23	82.1	420	8	ADT57937	Adt57937	Plant pol
259	23	82.1	281	3	ABG36021	Aag36021	Zea mays	332	23	82.1	423	8	ADT57937	Adt57937	Plant pol
260	23	82.1	291	3	ABG36020	Aag36020	Zea mays	333	23	82.1	425	4	AAW41711	Aaw41711	Human pol
261	23	82.1	303	2	AAW06422	Aaw06422	Yeast en-	334	23	82.1	426	8	ADJ49912	Adj49912	Oil-assoc
262	23	82.1	303	2	AAW06421	Aaw06421	Yeast en-	335	23	82.1	426	8	ADJ50126	Adj50126	Oil-assoc
263	23	82.1	303	2	AAW88102	Aaw88102	Transfera	336	23	82.1	426	8	ADS27875	Ads27875	Bacterial
264	23	82.1	303	3	AAV99490	Aay99490	Yeast acy	337	23	82.1	428	8	ADJ49280	Adj49280	Oil-assoc
265	23	82.1	303	3	ABE15286	Aae15286	Yeast LPA	338	23	82.1	432	8	ADN26876	Adn26876	Bacterial
266	23	82.1	303	6	ABR52952	Abr52952	Protein s	339	23	82.1	432	8	ADN26675	Adn26675	Bacterial
267	23	82.1	303	6	ABU24010	Abu24010	Protein e	340	23	82.1	433	3	AAW20957	Aaw20957	Arabidops
268	23	82.1	303	7	ADK62618	Adk62618	Disease t	341	23	82.1	436	6	ABU41319	Abu41319	Protein e
269	23	82.1	303	8	ADJ93787	Adj93787	Yeast lys	342	23	82.1	439	4	AAW93788	Aaw93788	Human pro
270	23	82.1	303	8	ADN49471	Adn49471	Yeast lys	343	23	82.1	439	8	ADY23217	Ady23217	Plant ful
271	23	82.1	303	8	ADU04610	Ado04610	Yeast lys	344	23	82.1	446	8	ADS28594	Ads28594	Bacterial
272	23	82.1	303	8	ADN18936	Adn18936	Bacterial	345	23	82.1	448	4	AAW82323	Aaw82323	S. epider
273	23	82.1	308	3	ABE43188	Aab43188	Human ORF	346	23	82.1	460	4	AAW78655	Aaw78655	Human pro
274	23	82.1	315	5	ADL16927	Adl16927	Murine NO	347	23	82.1	460	5	AAW68506	Aaw68506	Human GRP
275	23	82.1	315	7	ABW02144	Abw02144	Human GPC	348	23	82.1	468	5	ABP43604	Abp43604	Clone MGC
276	23	82.1	315	8	ADM42847	Adm42847	Murine od	349	23	82.1	469	8	ADT59033	Adt59033	Plant pol
277	23	82.1	319	8	ABE68087	AbE68087	Drosophil	350	23	82.1	470	4	AAW79639	Aaw79639	Human pro
278	23	82.1	319	9	ABE41680	Aeb41680	L. pneumo	351	23	82.1	470	6	ABU34745	Abu34745	Protein e
279	23	82.1	321	4	ABE61436	Abb61436	Drosophil	352	23	82.1	470	8	ADH22542	Adh22542	Human tra
280	23	82.1	321	4	ABE64718	Abb64718	Human sec	353	23	82.1	471	3	AAV90287	Aay90287	Human pep
281	23	82.1	321	4	ABE64717	Aab64717	Gene 5 hu	354	23	82.1	471	5	AAW47764	Aaw47764	Human pro
282	23	82.1	321	4	ABE65978	Aab65978	Human sec	355	23	82.1	471	5	AAW47764	Aaw47764	F-casett
283	23	82.1	321	4	ABE65977	Aab65977	Human sec	356	23	82.1	474	9	ABE39197	Aeb39197	L. pneumo
284	23	82.1	328	9	ABE38409	Aeb38409	L. pneumo	357	23	82.1	481	9	ABE35747	Aeb35747	L. pneumo
285	23	82.1	337	9	AEA99948	Aea99948	L. helvet	358	23	82.1	488	7	ABO66647	AbO66647	Klebsiell
286	23	82.1	344	4	ABG72619	Aag72619	Murine OR	359	23	82.1	491	6	ABU26104	Abu26104	Protein e
287	23	82.1	345	3	AAW42984	Aaw42984	Human ORF	360	23	82.1	502	6	ABU26099	Abu26099	Protein e
288	23	82.1	349	9	ADM26685	Adw26685	Fructo-ol	361	23	82.1	503	2	ABP57062	Abp57062	Rat chole
289	23	82.1	350	4	ABG94228	Aab94228	Human pro	362	23	82.1	503	6	ABP57062	Abp57062	Rat CYP7A
290	23	82.1	350	4	ABG08347	Abg08347	Novel hum	363	23	82.1	503	7	ADH48630	Adh48630	Rat prote
291	23	82.1	353	8	ADP67803	Adp67803	Novel hum	364	23	82.1	504	2	AAW72509	Aaw72509	Rat CYP7A
292	23	82.1	355	5	ABP66105	Abp66105	Bifidobac	365	23	82.1	504	2	AAW72509	Aaw72509	Rat CYP7A
293	23	82.1	357	5	ABP63001	Abp63001	Human pol	366	23	82.1	515	4	AAU33948	Aau33948	Hamster c
294	23	82.1	360	5	ABF41132	Abp41132	Human ova	367	23	82.1	523	4	ABE65104	AbE65104	Drosophil
295	23	82.1	368	4	ABG28311	Abg28311	Novel hum	368	23	82.1	525	7	ADE71266	AdE71266	Novel hum
296	23	82.1	370	4	ABE94432	Aab94432	Human pro	369	23	82.1	534	6	AAE30009	Aae30009	Tobacco c
297	23	82.1	378	7	ADK66908	Adk66908	Human toc	370	23	82.1	537	6	AAE30008	Aae30008	Tobacco c
298	23	82.1	383	2	AAW42447	Aaw42447	C. hetero	371	23	82.1	544	8	ADM97949	Adm97949	Protein e
299	23	82.1	385	4	ABE67960	Abbe67960	Drosophil	372	23	82.1	544	8	ADM97949	Adm97949	GFPp8D1
300	23	82.1	389	4	AAU34818	Aau34818	E. coli c	373	23	82.1	548	9	ADM97950	Adm97950	Valencia
301	23	82.1	389	8	ADK64184	Adk64184	Escherich	374	23	82.1	548	9	ADM97950	Adm97950	GFPp8D1
302	23	82.1	389	8	ADK13790	Adk13790	E. coli i	375	23	82.1	555	8	ADM97948	Adm97948	GFPp8B
303	23	82.1	397	4	ADL16915	Adl16915	Novel hum	376	23	82.1	557	9	ABU15931	Abu15931	Protein e
304	23	82.1	400	5	ADL16915	Adl16915	Rat NOVX	377	23	82.1	557	9	ABU15931	Abu15931	Protein e
305	23	82.1	403	4	AAE67025	Aag67025	Human SPF	378	23	82.1	566	2	AAW82635	Aaw82635	Amlyase f
306	23	82.1	403	4	AAE67024	Aag67024	Rat SPF	379	23	82.1	566	2	AAW82635	Aaw82635	Amlyase f
307	23	82.1	403	5	ABG61931	Abg61931	Prostate	380	23	82.1	566	2	AAW82635	Aaw82635	Amlyase f
308	23	82.1	403	5	ABE41422	Abp41422	Human ova	381	23	82.1	571	6	ABU43025	Abu43025	Protein e
309	23	82.1	403	5	ADL16913	Adl16913	Rat NOVX	382	23	82.1	571	6	ABU43025	Abu43025	Protein e
310	23	82.1	403	5	ADL16912	Adl16912	Human NOV	383	23	82.1	573	5	ABP40097	Abp40097	Staphyloc
311	23	82.1	403	5	ADL16914	Adl16914	Murine NO	384	23	82.1	573	5	ABP40097	Abp40097	Staphyloc
312	23	82.1	403	7	ADK66910	Adk66910	Human toc	385	23	82.1	574	4	AAU37358	Aau37358	Staphyloc
313	23	82.1	403	7	ABO66445	AbO66445	Klebsiell	386	23	82.1	574	4	AAU37358	Aau37358	Staphyloc
314	23	82.1	403	8	ADJ75589	Adj75589	Marker ge	387	23	82.1	574	6	ABU18989	Abu18989	Pathogen
315	23	82.1	403	8	ADN02622	Adn02622	Liver dis	388	23	82.1	574	6	ABW73119	Abw73119	Staphyloc
316	23	82.1	403	8	ADQ19056	Adq19056	Human sof	389	23	82.1	584	7	ADC94937	Adc94937	E. faeciu

390	23	82.1	610	8	ADN20025	Bacterial	Adn20025	463	23	82.1	896	8	ADQ89162	Human uro
391	23	82.1	615	7	Adf07764	Bacterial	Adf07764	464	23	82.1	896	3	ADR44918	Polypepti
392	23	82.1	617	7	AbO67804	Klebsiell	AbO67804	465	23	82.1	900	3	Aag47642	Arabidops
393	23	82.1	619	8	Adn27232	Bacterial	Adn27232	466	23	82.1	908	4	ABB65159	Drosophil
394	23	82.1	622	6	AdB06956	Alloiooc	AdB06956	467	23	82.1	916	7	ADI21204	Novel hum
395	23	82.1	633	4	ABG28306	Novel hum	ABG28306	468	23	82.1	941	9	ABM91757	M. xanthu
396	23	82.1	638	7	ABO69211	Pseudomon	ABO69211	469	23	82.1	948	3	AAG47641	Arabidops
397	23	82.1	640	8	AdR86494	Aspergill	AdR86494	470	23	82.1	975	5	ABG96290	Human ova
398	23	82.1	641	6	ABM68514	Photorhab	ABM68514	471	23	82.1	992	5	ABP29542	Streptoco
399	23	82.1	641	6	ABJ26298	Aspergill	ABJ26298	472	23	82.1	992	6	ABU46823	Protein e
400	23	82.1	645	2	Aaw22780	human RAD	Aaw22780	473	23	82.1	1002	8	ADR66344	Human pro
401	23	82.1	645	2	Aaw71293	Protein G	Aaw71293	474	23	82.1	1002	8	ADR66686	Human pro
402	23	82.1	652	4	ABB60647	Drosophil	ABB60647	475	23	82.1	1008	8	ADN22976	Bacterial
403	23	82.1	659	7	AdO4463	Bacterial	AdO4463	476	23	82.1	1008	8	ADN22977	Bacterial
404	23	82.1	670	6	ABM67836	Photorhab	ABM67836	477	23	82.1	1032	5	ABP27824	Streptoco
405	23	82.1	674	8	ADN2628	Bacterial	ADN2628	478	23	82.1	1053	8	ABO84970	Murine ca
406	23	82.1	675	6	ABJ25698	Aspergill	ABJ25698	479	23	82.1	1071	8	ADS43853	Bacterial
407	23	82.1	677	7	ABO69356	Pseudomon	ABO69356	480	23	82.1	1091	8	ADU24077	Human cys
408	23	82.1	679	8	ADN2629	Bacterial	ADN2629	481	23	82.1	1092	4	AAM79135	Human pro
409	23	82.1	694	8	ADN25736	Bacterial	ADN25736	482	23	82.1	1092	4	AAM40100	Human pol
410	23	82.1	695	3	RAY45097	Arabidops	AY45097	483	23	82.1	1092	8	ABO84971	Human can
411	23	82.1	695	6	AAO15964	Arabidops	AO15964	484	23	82.1	1094	4	AAO80119	Human pro
412	23	82.1	695	6	AAO15963	Arabidops	AO15963	485	23	82.1	1094	4	AAO41886	Human pol
413	23	82.1	695	9	ADV66378	Arabidops	ADV66378	486	23	82.1	1127	3	AAG47640	Arabidops
414	23	82.1	697	8	ADX94784	Plant ful	ADX94784	487	23	82.1	1173	9	ADY19886	PRO poly
415	23	82.1	707	4	RAM23744	Human EST	AM23744	488	23	82.1	1186	8	ADQ67807	Novel hum
416	23	82.1	707	4	RAM23715	Human EST	AM23715	489	23	82.1	1205	6	AAE36060	Human tra
417	23	82.1	719	6	ABU10383	Mouse pha	BU10383	490	23	82.1	1207	6	AAE36061	Human tra
418	23	82.1	720	5	ABP65117	Hypoxia-r	BP65117	491	23	82.1	1312	2	AAW22775	Human RAD
419	23	82.1	720	5	ABP59019	Human S4	BP59019	492	23	82.1	1312	2	AAW71295	Human hom
420	23	82.1	720	7	ADN13523	Human NOV	ADN13523	493	23	82.1	1312	9	ADY15954	PRO poly
421	23	82.1	720	9	ADZ84580	Engulfmen	ADZ84580	494	23	82.1	1312	9	ADY91604	Human pro
422	23	82.1	721	6	ABU10380	Human pha	BU10380	495	23	82.1	1318	5	ABY77985	Amino aci
423	23	82.1	739	8	ADS45167	Bacterial	ADS45167	496	23	82.1	1318	7	ADJ68860	Human hea
424	23	82.1	749	5	ABB04872	LDL recepr	BB04872	497	23	82.1	1318	8	ADJ66509	RAU50 hom
425	23	82.1	761	5	AAE13544	Human cer	AE13544	498	23	82.1	1318	8	ADS88326	Human pro
426	23	82.1	776	6	ADB12050	Alloiooc	Adb12050	499	23	82.1	1362	8	ADS34522	POSH prot
427	23	82.1	780	6	AAE34441	Human lip	AE34441	500	23	82.1	1394	8	ADN25124	Bacterial
428	23	82.1	794	8	ABO64392	Klebsiell	AO64392	501	23	82.1	1478	6	ABU35753	Protein e
429	23	82.1	815	8	ABM84044	Human dia	AM84044	502	23	82.1	1661	6	ABR39999	Human ABC
430	23	82.1	818	8	ADX73834	Plant ful	Adx73834	503	23	82.1	1864	2	AAW22602	Tyactome
431	23	82.1	820	7	ADN95393	Human BEC	ADN95393	504	23	82.1	1868	8	ADX56097	Streptomy
432	23	82.1	821	6	ABR92110	Human cer	ABr92110	505	23	82.1	1881	6	ABP98810	Human str
433	23	82.1	821	6	ABR92110	Human cer	ABr92110	506	23	82.1	2147	6	AAO31009	Human tra
434	23	82.1	821	7	ADE62260	Human pro	Ad62260	507	23	82.1	2654	8	ADK15820	Human ABC
435	23	82.1	821	7	ADA47954	Human pro	Ad47954	508	23	82.1	2660	8	ADK15816	Human ABC
436	23	82.1	821	8	ADL76337	Novel hum	AdL76337	509	23	82.1	2760	8	ADK15810	Human ABC
437	23	82.1	821	8	ADL70589	Cervical	AdL70589	510	23	82.1	3264	9	ADY72594	A. orient
438	23	82.1	821	8	ABM80342	Tumour-as	ABm80342	511	23	82.1	4952	8	ADK15818	Human ABC
439	23	82.1	821	8	ADP54308	Human pro	ADp54308	512	23	82.1	4958	8	ADK15793	Human ABC
440	23	82.1	821	8	ADU08455	Novel bro	ADU08455	513	23	82.1	5058	8	ADK15793	Human ABC
441	23	82.1	821	9	ADY14695	PRO poly	ADy14695	514	22	78.6	10	2	RAY22359	TPO recep
442	23	82.1	824	3	RAY95293	Human GEF	AY95293	515	22	78.6	10	3	AAI6983	TPO-mimet
443	23	82.1	828	4	ADS42481	Bacterial	ADS42481	516	22	78.6	10	5	ABY72869	TPO mimet
444	23	82.1	839	8	ADN11329	Aphis gos	ADn11329	517	22	78.6	10	7	ADJ73020	TPO mimet
445	23	82.1	843	7	ABO83863	Pseudomon	ABO83863	518	22	78.6	10	8	ADJ52655	CH1 delet
446	23	82.1	847	4	ABG17278	Novel hum	ABg17278	519	22	78.6	10	8	ADJ51616	CH1 delet
447	23	82.1	851	8	ADR10252	Human pro	ADR10252	520	22	78.6	15	2	RAY22387	TPO recep
448	23	82.1	855	4	AAW78466	Human pro	AAW78466	521	22	78.6	15	5	ABP57970	Human CD4
449	23	82.1	855	4	AAW38741	Human pol	AAW38741	522	22	78.6	20	9	ADZ98167	Human ami
450	23	82.1	855	7	ADE59602	Human pro	AdE59602	523	22	78.6	30	3	AAI38296	Human sec
451	23	82.1	855	8	ABM81053	Tumour-as	ABm81053	524	22	78.6	30	3	AAI38294	Human sec
452	23	82.1	868	7	ABM89313	Rice abio	ABm89313	525	22	78.6	32	4	AAO4838	Human pol
453	23	82.1	870	4	ABG30208	Novel hum	ABg30208	526	22	78.6	41	9	ABM96631	M. xanthu
454	23	82.1	871	8	ADY06939	Plant ful	ADy06939	527	22	78.6	46	3	AAI44949	Human thr
455	23	82.1	873	4	AAW79450	Human pro	AAW79450	528	22	78.6	46	3	RAY44948	Mouse thr
456	23	82.1	873	4	AAW40527	Human pol	AAW40527	529	22	78.6	48	4	AAO5200	Human pol
457	23	82.1	873	4	ASG17472	Novel hum	ABg17472	530	22	78.6	50	9	ADZ98172	Human ami
458	23	82.1	874	8	ADY24764	Plant ful	ADy24764	531	22	78.6	53	4	RAM88941	Human imm
459	23	82.1	874	8	ADS24763	Plant ful	ADy24763	532	22	78.6	60	7	ABR82695	Lbxi home
460	23	82.1	882	8	ADS21711	Bacterial	ADs21711	533	22	78.6	69	4	AAU51338	Propionib
461	23	82.1	896	6	ABR43913	Human sod	ABr43913	534	22	78.6	69	6	ABM47857	Propionib
462	23	82.1	896	7	ADJ68814	Human hea	ADj68814	535	22	78.6	70	5	ABP40467	Staphyloc

536	22	78.6	70	8	ADS06011	Staphyloc	609	22	78.6	223	3	AAG08497	Arabidops
537	22	78.6	72	3	AAG57925	Zea mays	610	22	78.6	224	3	AAG51820	Arabidops
538	22	78.6	73	7	ADG30700	Xanthomon	611	22	78.6	225	9	AEA10738	FGF-5 bin
539	22	78.6	74	7	AAO10156	Human pol	612	22	78.6	226	3	AAG49179	Arabidops
540	22	78.6	77	4	AAU40924	Propionib	613	22	78.6	227	3	AAG08496	Arabidops
541	22	78.6	87	4	AAU65713	Propionib	614	22	78.6	227	8	ADT56060	Plant pol
542	22	78.6	87	6	ABM37443	Propionib	615	22	78.6	228	8	ADF73127	Xenopus c
543	22	78.6	87	6	ABM62232	Propionib	616	22	78.6	228	8	ADM99656	African c
544	22	78.6	87	8	ADR95435	Novel S.	617	22	78.6	231	6	ABU29713	Protein e
545	22	78.6	92	8	ADR95435	Novel S.	618	22	78.6	232	6	ABU02611	S. pneumo
546	22	78.6	92	9	AEA59305	Streptoco	619	22	78.6	232	6	ABP81483	Streptoco
547	22	78.6	95	8	ADN47583	Thermoco	620	22	78.6	232	6	ABU44309	Protein e
548	22	78.6	95	4	ABG06872	Novel hum	621	22	78.6	232	6	ABU46283	Protein e
549	22	78.6	97	5	ABP42958	Human ova	622	22	78.6	232	8	ADK47700	Streptoco
550	22	78.6	99	7	ABO60953	Klebsiell	623	22	78.6	234	9	AEA49217	L. rhanno
551	22	78.6	100	3	AAG57924	Zea mays	624	22	78.6	235	3	AAG33167	Zea mays
552	22	78.6	100	5	ABU05471	M. tuberc	625	22	78.6	236	5	ABBA8744	Listeria
553	22	78.6	100	5	ADBT74329	Mycobacte	626	22	78.6	236	6	ABU32846	Protein e
554	22	78.6	101	7	ADB74435	Mycobacte	627	22	78.6	239	6	ABU38494	Protein e
555	22	78.6	105	3	AAV44960	KDEL rece	628	22	78.6	240	4	ABBS63201	Drosophil
556	22	78.6	109	3	AAV44961	KDEL rece	629	22	78.6	244	7	ADH87382	Enterococ
557	22	78.6	109	3	AAV44965	KDEL rece	630	22	78.6	244	7	ADH87382	Pseudomon
558	22	78.6	114	3	ABM85732	Mouse pro	631	22	78.6	252	6	ABU35766	Protein e
559	22	78.6	116	5	ABP08252	Human ORF	632	22	78.6	256	5	ABBS53587	Lactococc
560	22	78.6	116	6	ABP78249	N. gonorr	633	22	78.6	257	4	AAU87890	Propionib
561	22	78.6	117	8	ADP90619	Micromono	634	22	78.6	257	6	AAM64409	Propionib
562	22	78.6	120	8	ADX94476	Plant ful	636	22	78.6	266	2	AAW26705	Human hip
563	22	78.6	120	8	ADX90986	Plant ful	637	22	78.6	267	2	ADC96056	Human Cyp
564	22	78.6	125	5	ABP57969	Human CD4	638	22	78.6	267	8	ADK52135	Mouse ato
565	22	78.6	125	4	AAO03224	Human pol	639	22	78.6	270	6	ABU23259	Protein e
566	22	78.6	134	4	AAG82640	S. epider	640	22	78.6	271	7	ADL72407	Maize lon
567	22	78.6	136	3	AAG33199	Zea mays	641	22	78.6	277	3	AAG33166	Zea mays
568	22	78.6	136	3	AAG26815	Zea mays	642	22	78.6	277	9	ADZ17581	Theileria
569	22	78.6	136	4	AAO00478	Human pol	643	22	78.6	278	5	AAE16769	Human tra
570	22	78.6	137	4	ABB67359	Drosophil	644	22	78.6	278	8	ADU63798	A. oryzae
571	22	78.6	138	3	AAG33233	Zea mays	645	22	78.6	278	8	ADU63798	A. oryzae
572	22	78.6	138	8	ADY05350	Plant ful	646	22	78.6	280	9	ADM23376	Aspergill
573	22	78.6	141	3	AAB40305	Human ORF	647	22	78.6	291	3	AAG33165	Aspergill
574	22	78.6	141	5	ABP02707	Human ORF	648	22	78.6	294	6	ABP78112	Protein e
575	22	78.6	147	5	ABP38049	Staphyloc	649	22	78.6	294	8	ABU37516	Protein e
576	22	78.6	147	8	ADS07409	Staphyloc	650	22	78.6	294	8	ADU05985	M. catar
577	22	78.6	149	4	AAU37961	Streptoco	651	22	78.6	294	8	ADU05985	M. catar
578	22	78.6	153	8	ADV82718	Streptoco	652	22	78.6	294	8	ADT57068	Plant pol
579	22	78.6	154	6	ABU48752	Protein e	653	22	78.6	298	8	ADF45518	Pan trogl
580	22	78.6	159	6	AAE34015	Lithium pe	654	22	78.6	298	8	ADP29619	Human sec
581	22	78.6	160	6	ABU01398	S. pneumo	655	22	78.6	298	8	ADC28636	Bacterial
582	22	78.6	160	8	ADK48691	Streptoco	656	22	78.6	307	7	ADS95625	E. faeciu
583	22	78.6	160	8	ADK48691	Streptoco	657	22	78.6	307	7	ADS95625	E. faeciu
584	22	78.6	161	3	AAG15656	Arabidops	658	22	78.6	309	6	ADB10430	S. cerevis
585	22	78.6	161	4	ABBS2923	Escherich	659	22	78.6	310	4	AAE04582	Human olf
586	22	78.6	161	4	ABBS2957	Escherich	660	22	78.6	310	4	AAE04582	Human olf
587	22	78.6	161	8	ADR94930	Novel S.	661	22	78.6	310	5	AAU83574	Human nov
588	22	78.6	161	9	AEA58800	Streptoco	662	22	78.6	310	5	AAU83574	Human nov
589	22	78.6	173	3	AAG33198	Zea mays	663	22	78.6	310	5	AAU83574	Human nov
590	22	78.6	175	9	ADX02798	Sendai vi	664	22	78.6	310	7	ADC86081	Human GPC
591	22	78.6	176	8	ADR75253	Mouse gas	665	22	78.6	310	7	ADM29630	Novel hum
592	22	78.6	181	9	ADX02797	Sendai vi	666	22	78.6	310	7	ADM29630	Novel hum
593	22	78.6	182	2	AAV35544	C. pneumo	667	22	78.6	310	7	ADM29630	Novel hum
594	22	78.6	200	4	AAU54511	Propionib	668	22	78.6	314	4	ABBS59817	Drosophil
595	22	78.6	200	6	ABM51030	Propionib	669	22	78.6	315	3	AAG32359	Arabidops
596	22	78.6	203	1	AAV70394	Sequence	670	22	78.6	316	5	ADL16828	Murine NO
597	22	78.6	204	7	ADC85520	Sendai vi	671	22	78.6	316	5	ADL16828	Murine NO
598	22	78.6	204	8	ADX02796	Bacterial	672	22	78.6	316	7	ABW02145	Human GPC
599	22	78.6	205	8	ADS24392	Novel hum	673	22	78.6	316	7	ABW02145	Human GPC
600	22	78.6	214	4	AGT72857	Human olf	674	22	78.6	316	9	ADQ98158	Mouse olf
601	22	78.6	215	9	ADX02795	Sendai vi	675	22	78.6	317	3	AAZ139245	Wheat pan
602	22	78.6	218	4	AG90916	C glutami	676	22	78.6	318	3	AAZ139245	Wheat pan
603	22	78.6	218	9	ABE15287	C glutami	677	22	78.6	318	3	AAZ139245	Wheat pan
604	22	78.6	221	5	AAG32360	Arabidops	678	22	78.6	319	8	ADY08458	Plant ful
605	22	78.6	222	5	ABB48132	Listeria	679	22	78.6	320	7	ADF06693	Bacterial
606	22	78.6	222	6	ABU32818	Protein e	680	22	78.6	322	4	AAU47175	Propionib
607	22	78.6	223	3	AAZ17048	Arabidops	681	22	78.6	322	6	ABM70262	Photorhab
608	22	78.6	223	3	AAG49180	Arabidops							

682	22	78.6	322	6	ABM43694	Abm43694 Propionib	755	22	78.6	388	5	ABB93126	Abb93126 Herbicida
683	22	78.6	324	4	ABB70055	Abb70055 Drosophil	756	22	78.6	388	8	ADS24489	Ads24489 Bacterial
684	22	78.6	326	4	AAg72856	Aag72856 Human olf	757	22	78.6	389	3	AAg51818	Aag51818 Arabidops
685	22	78.6	328	3	AAg06187	Aag06187 Arabidops	758	22	78.6	391	3	AAg06185	Aag06185 Arabidops
686	22	78.6	328	3	AAg49302	Aag49302 Arabidops	759	22	78.6	391	3	AAg49300	Aag49300 Arabidops
687	22	78.6	328	8	ADx70908	Adx70908 Plant ful	760	22	78.6	391	5	ABB93709	Abb93709 Herbicida
688	22	78.6	330	6	ABU21137	Abu21137 Protein e	761	22	78.6	391	6	ABU47379	Abu47379 Protein e
689	22	78.6	331	5	ABP73685	Abp73685 Candida a	762	22	78.6	392	8	ADY05588	Ady05588 Plant ful
690	22	78.6	332	4	ABG28228	Abg28228 Novel hum	763	22	78.6	392	4	AAU38299	Aau38299 Salmonell
691	22	78.6	332	6	ABU49437	Abu49437 Protein e	764	22	78.6	398	8	ADS22029	Ads22029 Bacterial
692	22	78.6	333	6	ABU47009	Abu47009 Protein e	765	22	78.6	398	8	ADx74065	Adx74065 Plant ful
693	22	78.6	333	8	ADQ37063	Adq37063 Cell prol	766	22	78.6	398	8	ADx92502	Adx92502 Plant ful
694	22	78.6	337	3	AAg51653	Aag51653 Arabidops	767	22	78.6	399	2	AAr74660	Aar74660 Pseudomon
695	22	78.6	340	9	AEA79823	Aea79823 BCs103 ce	768	22	78.6	400	2	AAW22521	Aaw22521 Bacillus
696	22	78.6	342	6	ABM65247	Abm65247 Photorhab	769	22	78.6	400	2	AAW23601	Aaw23601 Bacillus
697	22	78.6	343	3	AAAB32510	Abm32510 S. lavend	770	22	78.6	400	2	AAW57431	Aaw57431 Bacillus
698	22	78.6	343	7	ADE10282	Adel0282 S. lavend	771	22	78.6	401	6	ABU19990	Abu19990 Protein e
699	22	78.6	344	3	AAg10368	Aag10368 Arabidops	772	22	78.6	403	6	ABU49552	Abu49552 Protein e
700	22	78.6	344	3	AAg45683	Aag45683 Arabidops	773	22	78.6	404	6	ADb07808	Adb07808 Alloiococ
701	22	78.6	345	6	ABU17023	Abu17023 Protein e	774	22	78.6	405	8	ADY04935	Ady04935 Plant ful
702	22	78.6	346	4	ABG02100	Abg02100 Novel hum	775	22	78.6	407	6	ABU40512	Abu40512 Protein e
703	22	78.6	347	4	AAg72620	Aag72620 Murine OR	776	22	78.6	407	8	ADx93230	Adx93230 Plant ful
704	22	78.6	349	6	ABU44697	Abu44697 Protein e	777	22	78.6	409	1	AAp81843	Aap81843 Sequence
705	22	78.6	350	8	ADA33389	Ada33389 Acinetoba	778	22	78.6	410	2	AAW12378	Aaw12378 P300-CelB
706	22	78.6	350	8	ADS43305	Ads43305 Bacterial	779	22	78.6	410	6	ADb12347	Adb12347 Alloiococ
707	22	78.6	355	4	ABG28230	Abg28230 Novel hum	780	22	78.6	411	2	AAW12379	Aaw12379 P300-CelB
708	22	78.6	356	6	ABM68580	Abm68580 Photorhab	781	22	78.6	411	2	AAW12381	Aaw12381 P300-CelB
709	22	78.6	356	6	ABM69316	Abm69316 Photorhab	782	22	78.6	412	2	AAW12380	Aaw12380 P300-CelB
710	22	78.6	356	6	ABM68997	Abm68997 Photorhab	783	22	78.6	412	6	ABU33108	Abu33108 Protein e
711	22	78.6	356	6	ABM67343	Abm67343 Photorhab	784	22	78.6	412	9	ABE41502	Aeb41502 L. pneumo
712	22	78.6	356	6	ABM67680	Abm67680 Photorhab	785	22	78.6	413	8	ADx77667	Adx77667 Plant ful
713	22	78.6	356	6	ABM68134	Abm68134 Photorhab	786	22	78.6	415	7	ABM89908	Abm89908 Rice abio
714	22	78.6	356	6	ABM68359	Abm68359 Photorhab	787	22	78.6	416	6	ABU30633	Abu30633 Protein e
715	22	78.6	358	7	AAU86427	AAu86427 Rice abio	788	22	78.6	417	8	ADN25102	Adn25102 Bacterial
716	22	78.6	360	4	AAU24527	AAu24527 Human olf	789	22	78.6	417	8	ADN25102	Adn25102 Bacterial
717	22	78.6	360	5	AAU85151	AAu85151 G-coupled	790	22	78.6	420	6	ABU39017	Abu39017 Protein e
718	22	78.6	366	8	ADN23745	Adn23745 Bacterial	791	22	78.6	420	3	AAg29585	Aag29585 Arabidops
719	22	78.6	369	4	AAU50912	AAu50912 Propionib	792	22	78.6	424	8	ADJ56879	Adj56879 Arabidops
720	22	78.6	369	6	ABM47431	Abm47431 Propionib	793	22	78.6	424	9	ABE38215	Aeb38215 L. pneumo
721	22	78.6	370	4	AAAB79723	AAb79723 Coryneb	794	22	78.6	425	3	AAg29584	Aag29584 Arabidops
722	22	78.6	370	4	AAAB79728	AAb79728 Coryneb	795	22	78.6	425	8	ADx79792	Adx79792 Plant ful
723	22	78.6	370	4	AAU71899	AAu71899 Coryneb	796	22	78.6	426	3	AAg13338	Aag13338 Arabidops
724	22	78.6	370	4	AAU71899	AAu71899 Coryneb	797	22	78.6	426	3	AAg51652	Aag51652 Arabidops
725	22	78.6	372	6	ADA34058	Ada34058 Acinetoba	798	22	78.6	428	9	AEA29532	Aea29532 Fusion pr
726	22	78.6	374	8	ADU47402	Adu47402 Coryneb	799	22	78.6	429	6	ABU47820	Abu47820 Protein e
727	22	78.6	377	4	AAAB79727	AAb79727 Coryneb	800	22	78.6	429	6	ABU47200	Abu47200 Protein e
728	22	78.6	377	4	AAU71898	AAu71898 Coryneb	801	22	78.6	430	3	AAg29583	Aag29583 Arabidops
729	22	78.6	377	4	AAU71898	AAu71898 Coryneb	802	22	78.6	431	6	ADb12085	Adb12085 Alloiococ
730	22	78.6	378	8	ADJ49740	Adj49740 Oil-assoc	803	22	78.6	432	6	ADB08864	Adb08864 Alloiococ
731	22	78.6	379	8	ADO80279	Ado80279 Mucobacte	804	22	78.6	433	6	ADb12345	Adb12345 Alloiococ
732	22	78.6	381	4	AAAB52461	AAb52461 Mycobacte	805	22	78.6	433	8	ADJ49755	Adj49755 Oil-assoc
733	22	78.6	381	7	ADC60855	Adc60855 M. tuberc	806	22	78.6	433	8	ADS22912	Ads22912 Bacterial
734	22	78.6	382	6	ADB07806	Adb07806 Alloiococ	807	22	78.6	435	3	AAW75574	Aaw75574 Neisseria
735	22	78.6	382	8	ADO80277	Ado80277 Mycobacte	808	22	78.6	440	6	ABM70591	Abm70591 Photorhab
736	22	78.6	385	2	AAI39314	Aay39314 SpuR prot	809	22	78.6	443	8	ADx91485	Adx91485 Plant ful
737	22	78.6	385	3	AAAG49301	Aag49301 Arabidops	810	22	78.6	444	8	ADS24875	Ads24875 Bacterial
738	22	78.6	385	3	AAAG06186	Aag06186 Arabidops	811	22	78.6	447	4	AAg74174	Aag74174 Human col
739	22	78.6	385	4	AAAB70961	AAb70961 S. spinos	812	22	78.6	447	8	ABU22101	Abu22101 Protein e
740	22	78.6	385	9	ADZ99205	Adz99205 S. spinos	813	22	78.6	447	8	ADJ49780	Adj49780 Oil-assoc
741	22	78.6	385	9	ADZ99291	Adz99291 S. spinos	814	22	78.6	448	8	ADY11816	Ady11816 Plant ful
742	22	78.6	385	9	ADZ99299	Adz99299 S. spinos	815	22	78.6	451	5	ABP69683	Abp69683 Human pol
743	22	78.6	385	9	AEA10735	Aea10735 FGF-5 bin	816	22	78.6	455	4	ABB11613	Abb11613 Human gor
744	22	78.6	385	9	AEA10742	Aea10742 FGF-5 bin	817	22	78.6	457	8	ADN25370	Adn25370 Bacterial
745	22	78.6	385	9	AEA10731	Aea10731 FGF-5 bin	818	22	78.6	460	7	ABO79022	AbO79022 Pseudomon
746	22	78.6	385	9	AEA10733	Aea10733 FGF-5 bin	819	22	78.6	461	5	ABB05724	Abb05724 Human tes
747	22	78.6	385	9	AEA10740	Aea10740 FGF-5 bin	820	22	78.6	462	2	AAW57433	Aaw57433 Cloned al
748	22	78.6	385	9	AEA10727	Aea10727 FGF-5 bin	821	22	78.6	467	2	AAW05731	Aaw05731 Cellulase
749	22	78.6	385	9	AEA10729	Aea10729 FGF-5 bin	822	22	78.6	467	2	AAW00382	Aaw00382 Bacillus
750	22	78.6	386	3	AAAG32358	Aag32358 Arabidops	823	22	78.6	467	5	ABB81990	Abb81990 Bacillus
751	22	78.6	387	6	ABU45056	Abu45056 Protein e	824	22	78.6	467	8	ABU17917	Abu17917 Protein e
752	22	78.6	387	7	ADY70100	Ady70100 Human hea	825	22	78.6	469	8	ADN60465	Adn60465 B. lichen
753	22	78.6	387	8	ADO62666	Ado62666 Transcrip	826	22	78.6	470	6	ABU17011	Abu17011 Protein e
754	22	78.6	388	3	AAg17046	Aag17046 Arabidops	827	22	78.6	471	9	ABE41489	Aeb41489 L. pneumo

828	22	78.6	472	6	ABU18104	Abu18104 Protein e	901	22	78.6	584	8	ADS28133	AdS28133 Bacterial
829	22	78.6	473	5	ABG32151	Arabidops	902	22	78.6	592	7	ABO64264	ABO64264 Klebsiell
830	22	78.6	474	3	ADY64606	S. mansoni	903	22	78.6	592	9	AEA43084	AEA43084 M. smegma
831	22	78.6	475	3	AAY75573	Neisseria	904	22	78.6	592	9	AEA43085	AEA43085 M. smegma
832	22	78.6	475	6	AAE36320	Moraxella	905	22	78.6	593	8	ADS14851	ADS14851 Pseudomon
833	22	78.6	476	6	AAE36320	Moraxella	906	22	78.6	595	7	ABO78285	ABO78285 Pseudomon
834	22	78.6	477	6	ADB10428	Alloioiococ	907	22	78.6	605	8	ADJ49335	ADJ49335 Oil-assoc
835	22	78.6	481	6	ADA33285	Actinetoba	908	22	78.6	616	7	ABO79391	ABO79391 Pseudomon
836	22	78.6	482	4	AAU38270	Salmonell	909	22	78.6	618	5	ABP26565	ABP26565 Streptoco
837	22	78.6	484	8	ADN18452	Bacterial	910	22	78.6	618	5	ABU46844	ABU46844 Protein e
838	22	78.6	485	8	ADU07677	Amino aci	911	22	78.6	619	7	ADH87640	ADH87640 Enterococ
839	22	78.6	485	9	ABE38200	L. pneumo	912	22	78.6	624	8	ADX97030	ADX97030 Plant ful
840	22	78.6	488	3	AGS11651	Arabidops	913	22	78.6	629	7	ABO77217	ABO77217 Pseudomon
841	22	78.6	488	3	AGS11651	Arabidops	914	22	78.6	636	5	AAW49536	AAW49536 S. cynthi
842	22	78.6	488	4	ABB60032	Drosophil	915	22	78.6	648	7	ABO71382	ABO71382 Pseudomon
843	22	78.6	488	5	ABG32153	Arabidops	916	22	78.6	652	5	ABP66325	ABP66325 Bifidobac
844	22	78.6	488	7	ADB95062	A. thalia	917	22	78.6	660	8	ADN22630	ADN22630 Bacterial
845	22	78.6	488	8	ADH38674	Aaribops	918	22	78.6	661	2	AAE26176	AAE26176 GORgab pr
846	22	78.6	488	8	ADH38674	Aaribops	919	22	78.6	661	2	AAE26176	AAE26176 GORgab pr
847	22	78.6	489	8	ADH87377	Novel hum	920	22	78.6	662	8	AAE22674	AAE22674 Recombina
848	22	78.6	495	4	ABG17680	Novel hum	921	22	78.6	662	8	AAE22674	AAE22674 Recombina
849	22	78.6	495	5	AAE22310	Methylomo	922	22	78.6	673	4	ABB59262	ABB59262 Drosophil
850	22	78.6	497	5	ABG61589	High grow	923	22	78.6	673	4	ABB59262	ABB59262 Drosophil
851	22	78.6	497	5	AAU80333	Methylomo	924	22	78.6	675	4	ABG16318	ABG16318 Novel hum
852	22	78.6	497	6	ADA14538	Methylomo	925	22	78.6	675	4	ABG16318	ABG16318 Novel hum
853	22	78.6	497	9	ADW08985	Methylomo	926	22	78.6	679	7	ABO79541	ABO79541 Pseudomon
854	22	78.6	497	9	ABN94498	M. xanthu	927	22	78.6	690	4	ABB61438	ABB61438 Drosophil
855	22	78.6	504	8	ADJ49680	Oil-assoc	928	22	78.6	690	4	ABB61438	ABB61438 Drosophil
856	22	78.6	504	9	ADJ49680	Oil-assoc	929	22	78.6	696	7	ABO71841	ABO71841 Pseudomon
857	22	78.6	511	8	ADY07023	Plant ful	930	22	78.6	699	8	ADI26806	ADI26806 Saccharom
858	22	78.6	514	5	AAU80014	Coniferal	931	22	78.6	709	6	ABU40849	ABU40849 Protein e
859	22	78.6	514	7	ADP93898	Quaking a	932	22	78.6	712	6	ABU25450	ABU25450 Aspergill
860	22	78.6	517	4	AAU60063	Propionib	933	22	78.6	716	8	ADL06147	ADL06147 M. catarr
861	22	78.6	520	8	ABM56582	Bacterial	934	22	78.6	716	8	ADL06147	ADL06147 M. catarr
862	22	78.6	521	8	ADS27891	Bacterial	935	22	78.6	719	7	ABM90154	ABM90154 Rice abio
863	22	78.6	522	8	ADN20865	Bacterial	936	22	78.6	720	7	ADP05453	ADP05453 Bacterial
864	22	78.6	525	9	ADZ46904	BASB232 p	937	22	78.6	724	6	ABU10378	ABU10378 Drosophil
865	22	78.6	526	8	ADS29015	Bacterial	938	22	78.6	725	8	ADS43821	ADS43821 Bacterial
866	22	78.6	527	3	AAU13336	Arabidops	939	22	78.6	736	6	ABU61998	ABU61998 Human thr
867	22	78.6	529	3	AAU13336	Arabidops	940	22	78.6	736	6	ABU61998	ABU61998 Human thr
868	22	78.6	533	2	AAU12276	Recombina	941	22	78.6	736	6	ABU61998	ABU61998 Human thr
869	22	78.6	533	2	AAU12276	Recombina	942	22	78.6	736	6	ABU61998	ABU61998 Human thr
870	22	78.6	533	4	AAU37798	Streptoco	943	22	78.6	755	7	ADK63620	ADK63620 Disease t
871	22	78.6	533	4	AAU37798	Streptoco	944	22	78.6	755	7	ADK63620	ADK63620 Disease t
872	22	78.6	533	5	ADP26773	Streptoco	945	22	78.6	755	7	ADK63620	ADK63620 Disease t
873	22	78.6	533	6	ABU00989	S. pneumo	946	22	78.6	755	7	ADK63620	ADK63620 Disease t
874	22	78.6	533	6	ABU45930	Protein e	947	22	78.6	755	7	ADK63620	ADK63620 Disease t
875	22	78.6	533	8	ADK48791	Streptoco	948	22	78.6	755	7	ADK63620	ADK63620 Disease t
876	22	78.6	533	8	ADT50098	Streptoco	949	22	78.6	755	7	ADK63620	ADK63620 Disease t
877	22	78.6	533	8	ADV89309	Streptoco	950	22	78.6	755	7	ADK63620	ADK63620 Disease t
878	22	78.6	533	8	ADV82733	Streptoco	951	22	78.6	755	7	ADK63620	ADK63620 Disease t
879	22	78.6	533	8	ADV80562	Streptoco	952	22	78.6	755	7	ADK63620	ADK63620 Disease t
880	22	78.6	534	6	ABU44198	Protein e	953	22	78.6	755	7	ADK63620	ADK63620 Disease t
881	22	78.6	535	5	ABBA7951	Listeria	954	22	78.6	755	7	ADK63620	ADK63620 Disease t
882	22	78.6	535	6	ABU22531	Protein e	955	22	78.6	755	7	ADK63620	ADK63620 Disease t
883	22	78.6	535	6	ADK15044	Pseudomon	956	22	78.6	755	7	ADK63620	ADK63620 Disease t
884	22	78.6	535	8	ADK15044	Pseudomon	957	22	78.6	755	7	ADK63620	ADK63620 Disease t
885	22	78.6	535	8	ADK15044	Pseudomon	958	22	78.6	755	7	ADK63620	ADK63620 Disease t
886	22	78.6	535	8	ADK15044	Pseudomon	959	22	78.6	755	7	ADK63620	ADK63620 Disease t
887	22	78.6	535	8	ADK15044	Pseudomon	960	22	78.6	755	7	ADK63620	ADK63620 Disease t
888	22	78.6	535	8	ADK15044	Pseudomon	961	22	78.6	755	7	ADK63620	ADK63620 Disease t
889	22	78.6	535	8	ADK15044	Pseudomon	962	22	78.6	755	7	ADK63620	ADK63620 Disease t
890	22	78.6	535	8	ADK15044	Pseudomon	963	22	78.6	755	7	ADK63620	ADK63620 Disease t
891	22	78.6	535	8	ADK15044	Pseudomon	964	22	78.6	755	7	ADK63620	ADK63620 Disease t
892	22	78.6	535	8	ADK15044	Pseudomon	965	22	78.6	755	7	ADK63620	ADK63620 Disease t
893	22	78.6	535	8	ADK15044	Pseudomon	966	22	78.6	755	7	ADK63620	ADK63620 Disease t
894	22	78.6	535	8	ADK15044	Pseudomon	967	22	78.6	755	7	ADK63620	ADK63620 Disease t
895	22	78.6	535	8	ADK15044	Pseudomon	968	22	78.6	755	7	ADK63620	ADK63620 Disease t
896	22	78.6	535	8	ADK15044	Pseudomon	969	22	78.6	755	7	ADK63620	ADK63620 Disease t
897	22	78.6	535	8	ADK15044	Pseudomon	970	22	78.6	755	7	ADK63620	ADK63620 Disease t
898	22	78.6	535	8	ADK15044	Pseudomon	971	22	78.6	755	7	ADK63620	ADK63620 Disease t
899	22	78.6	535	8	ADK15044	Pseudomon	972	22	78.6	755	7	ADK63620	ADK63620 Disease t
900	22	78.6	535	8	ADK15044	Pseudomon	973	22	78.6	755	7	ADK63620	ADK63620 Disease t

AdS28133	Bacterial
ABO64264	Klebsiell
AEA43084	M. smegna
AEA43085	M. smegna
ADS14851	Pseudomon
ABO78285	Pseudomon
ADJ49335	Oil-assoc
ABO79391	Pseudomon
ABP26565	Streptoco
ABU46844	Protein e
ADH87640	Enterococ
ADX97030	Plant ful
ABO77217	Pseudomon
AAW49536	S. cynthi
ABO71382	Pseudomon
ABP66325	Bifidobac
ADN22630	Bacterial
AAE26176	GORGab pr
AAE22674	Recombina
ABB59262	Drosophil
ABG79541	Pseudomon
ABG16318	Novel hum
ABG16244	Human the
ABO80141	Pseudomon
ABB66108	Drosophil
ABB61438	Drosophil
ABO71841	Pseudomon
ADI26806	Saccharom
ABU40849	Protein e
ABJ25450	Aspergill
ADL06147	M. catarr
ABM65573	Drosophil
ABM86452	Rice abio
ABM90154	Rice abio
ADP05453	Bacterial
ABU61429	Drosophil
ABU01378	Drosophil
ADG43821	Bacterial
ABU61998	Human thr
ADG86001	Aspergill
ADK63620	Disease t
ADN19332	Bacterial
AAW93335	Human pol
ADJ130836	Human pro
ADG80079	Aspergill
ADG84693	Rat Prote
ADL82106	Fibrinoge
ABM53413	Protein s
ADK64704	Disease t
ADG08397	Novel pro
ADG45519	Rat KIAA1
ABU05813	M. tuberc
ABU36595	Protein e
ABU34905	Protein e
ADY22311	Plant ful
ABU34793	Protein e
ABU36818	Protein e
ADK93197	Plant ful
ABE17427	E. coli K
ADN20859	Bacterial
ADK4572	Klebsiell
ADG42628	Bacterial
ABU33969	Protein e
ABM30297	Rice abio
ABM94038	L. pneumo
ABM84487	Human dia

974 22 78.6 949 8 ABM84486
 975 22 78.6 956 7 ADP65204
 976 22 78.6 956 8 ADQ19408
 977 22 78.6 957 8 ADS27052
 978 22 78.6 958 8 ABM84485
 979 22 78.6 960 2 AAY29797
 980 22 78.6 960 4 AAB50089
 981 22 78.6 960 5 ABB05633
 982 22 78.6 960 6 ABP70633
 983 22 78.6 960 6 ABM04786
 984 22 78.6 963 8 ABM84484
 985 22 78.6 968 8 ABM84484
 986 22 78.6 1031 9 ABM93606
 987 22 78.6 1044 8 ADT55557
 988 22 78.6 1052 7 ABM89794
 989 22 78.6 1056 3 AAG31894
 990 22 78.6 1073 3 AAG31893
 991 22 78.6 1079 3 AAG48607
 992 22 78.6 1087 3 AAG31892
 993 22 78.6 1087 7 ADB67081
 994 22 78.6 1087 8 ADF45529
 995 22 78.6 1087 8 ADL99352
 996 22 78.6 1096 3 AAG48606
 997 22 78.6 1104 8 ADN21482
 998 22 78.6 1110 3 AAG48605
 999 22 78.6 1120 7 ADJ72235
 1000 22 78.6 1238 4 ABB62022

ALIGNMENTS

RESULT 1
 AAY22365
 ID AAY22365 standard; peptide; 11 AA.

AC AAY22365;

XX 27-SEP-1999 (first entry)

XX TPO receptor binding peptide sequence, SEQ ID NO. 16.

XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX Synthetic.

XX US5932546-A.

XX 03-AUG-1999.

XX 04-OCT-1996; 96US-00726464.

XX 04-OCT-1996; 96US-00726464.

XX (GLAX) GLAXO WELLCOME INC.

XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

XX Cwirla SE, Johnson SS;

XX WPI; 1999-457122/38.

XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopaenia.

XX Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 CC mM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders of thrombocytopaenia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration

XX Sequence 11 AA;

Query Match 92.9%; Score 26; DB 2; Length 11;

Best Local Similarity 66.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

Db 3 VRDQIMLSL 11

RESULT 2

AAY22370

ID AAY22370 standard; peptide; 11 AA.

XX AC AAY22370;

XX 27-SEP-1999 (first entry)

XX TPO receptor binding peptide sequence, SEQ ID NO. 21.

XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX Synthetic.

XX US5932546-A.

XX 03-AUG-1999.

XX 04-OCT-1996; 96US-00726464.

XX 04-OCT-1996; 96US-00726464.

XX (GLAX) GLAXO WELLCOME INC.

XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

XX Cwirla SE, Johnson SS;

XX WPI; 1999-457122/38.

XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopaenia.

XX Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 CC mM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders of thrombocytopaenia,
 CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration

XX Sequence 11 AA;

Query Match 92.9%; Score 26; DB 2; Length 11;

Best Local Similarity 66.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9


```
Db          ||||| 3 VRDQIWAAL 11
RESULT 3
AAB16986
ID AAB16986 standard; peptide; 11 AA.
XX
AC AAB16986;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:42.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
XX
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 210; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linker; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
DB 3 VRDQIWAAL 11
RESULT 5
AAB72872
```

```
Db          ||||| 3 VRDQIWAAL 11
RESULT 3
AAB16986
ID AAB16986 standard; peptide; 11 AA.
XX
AC AAB16986;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:42.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
XX
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 210; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linker; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;
XX
Query Match 92.9%; Score 26; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
DB 3 VRDQIWAAL 11
```

ID ABB72872 standard; peptide; 11 AA.
 AC ABB72872;
 DT 05-APR-2002 (first entry)
 DE TPO mimetic peptide SEQ ID NO:42.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200183525-A2.
 PN 08-NOV-2001.
 XX
 XX 02-MAY-2001; 2001WO-US014310.
 XX 03-MAY-2000; 2000US-00563286.
 XX (AMGE-) AMGEN INC.
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX WPI; 2002-130313/17.
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 XX diabetic retinopathy, obesity, sleep disorders and infertility.
 XX Claim 39; Page 43; 176pp; English.
 XX The present invention describes a vehicle-peptide molecule (I) or its
 XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 XX cycostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 XX antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 XX neuroprotective activities. (I) can be used as a therapeutic or
 XX prophylactic agent as well as for screening purposes. (I) is useful for
 XX diagnosing diseases characterised by dysfunction of their associated
 XX protein of interest, for identifying normal or abnormal proteins of
 XX interest, as a part of diagnostic kit to detect the presence of their
 XX proteins of interest in a biological sample. Additionally, (I) is useful
 XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 XX infertility, and neurological degenerative diseases. (I), comprising EPO-
 XX mimetic compounds are useful for treating disorders characterised by low
 XX red blood cell levels such as anaemia. The TPO-mimetic comprising
 XX compounds are useful for treating conditions that involve an existing
 XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 XX deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 XX tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 XX and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 XX represent amino acid and nucleic acid sequences used in the
 XX exemplification of the present invention
 XX Sequence 11 AA;
 XX Query Match 92.9%; Score 26; DB 5; Length 11;
 XX Best Local Similarity 66.7%; Pred. No. 12;
 XX Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXL 9
 DB 3 VRDQIWAAL 11
 RESULT 6
 ID ABB72874
 XX ABB72874 standard; peptide; 11 AA.
 AC ABB72874;
 XX
 DT 05-APR-2002 (first entry)
 DE TPO mimetic peptide SEQ ID NO:44.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200183525-A2.
 PN 08-NOV-2001.
 XX
 XX 02-MAY-2001; 2001WO-US014310.
 XX 03-MAY-2000; 2000US-00563286.
 XX (AMGE-) AMGEN INC.
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX WPI; 2002-130313/17.
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 XX diabetic retinopathy, obesity, sleep disorders and infertility.
 XX Claim 39; Page 43; 176pp; English.
 XX The present invention describes a vehicle-peptide molecule (I) or its
 XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 XX cycostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 XX antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 XX neuroprotective activities. (I) can be used as a therapeutic or
 XX prophylactic agent as well as for screening purposes. (I) is useful for
 XX diagnosing diseases characterised by dysfunction of their associated
 XX protein of interest, for identifying normal or abnormal proteins of
 XX interest, as a part of diagnostic kit to detect the presence of their
 XX proteins of interest in a biological sample. Additionally, (I) is useful
 XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 XX infertility, and neurological degenerative diseases. (I), comprising EPO-
 XX mimetic compounds are useful for treating disorders characterised by low
 XX red blood cell levels such as anaemia. The TPO-mimetic comprising
 XX compounds are useful for treating conditions that involve an existing
 XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 XX deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 XX tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 XX and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 XX represent amino acid and nucleic acid sequences used in the
 XX exemplification of the present invention

XX SQ Sequence 11 AA;
 Query Match 92.9%; Score 26; DB 5; Length 11;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 7
 ADJ73023
 ID ADJ73023 standard; peptide; 11 AA.
 XX AC ADJ73023;
 XX DT 06-MAY-2004 (first entry)
 XX DE TPO mimetic peptide sequence SeqID 477.
 XX KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 XX KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 XX KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 XX KW TPO.
 XX OS Synthetic.
 XX PN WO2003084477-A2.
 XX PD 16-OCT-2003.
 XX PF 24-MAR-2003; 2003WO-US009139.
 XX PR 29-MAR-2002; 2002US-0368791P.
 XX PA (CENZ) CENTOCOR INC.
 XX PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
 XX DR WPI; 2003-804237/75.
 XX PT New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX PS Disclosure; SEQ ID NO 477; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is a TPO mimetic peptide sequence used to make a
 CC mimetibody of the invention.

XX SQ Sequence 11 AA;
 Query Match 92.9%; Score 26; DB 7; Length 11;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 9
 ADJ52658

Oy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 8
 ADJ73025
 ID ADJ73025 standard; peptide; 11 AA.
 XX AC ADJ73025;
 XX DT 06-MAY-2004 (first entry)
 XX DE TPO mimetic peptide sequence SeqID 479.
 XX KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 XX KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 XX KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 XX KW TPO.
 XX OS Synthetic.
 XX PN WO2003084477-A2.
 XX PD 16-OCT-2003.
 XX PF 24-MAR-2003; 2003WO-US009139.
 XX PR 29-MAR-2002; 2002US-0368791P.
 XX PA (CENZ) CENTOCOR INC.
 XX PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
 XX DR WPI; 2003-804237/75.
 XX PT New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX PS Disclosure; SEQ ID NO 479; 97pp; English.

XX CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is a TPO mimetic peptide sequence used to make a
 CC mimetibody of the invention.

XX SQ Sequence 11 AA;
 Query Match 92.9%; Score 26; DB 7; Length 11;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 9
 ADJ52658

```

ID AC XX XX ADJ52658 standard; peptide; 11 AA.
AC ADJ52658;
XX DT XX 06-MAY-2004 (first entry)
XX DE XX CHI deleted mimetibody-related peptide SeqID477.
XX XX
XX KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX OS Unidentified.
XX OS Synthetic.
XX PN WO2004002417-A2.
XX PD 08-JAN-2004.
XX PF 27-JUN-2003; 2003WO-US020347.
XX PR 28-JUN-2002; 2002US-0392431P.
XX PA (CENZ ) CENTOCOR INC.
XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX DR WPI; 2004-082870/08.
XX PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX PS Claim 2; SEQ ID NO 477; 129pp; English.
XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX SQ Sequence 11 AA;
XX Query Match 92.9%; Score 26; DB 8; Length 11;
XX Best Local Similarity 66.7%; Pred. No. 12;
XX Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB |||||
3 VRDQIWAAL 11

RESULT 10
ADJ52660
ID ADJ52660 standard; peptide; 11 AA.
XX AC ADJ52660;

ID AC XX XX ADJ52658 standard; peptide; 11 AA.
AC ADJ52658;
XX DT XX 06-MAY-2004 (first entry)
XX DE XX CHI deleted mimetibody-related peptide SeqID477.
XX XX
XX KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX OS Unidentified.
XX OS Synthetic.
XX PN WO2004002417-A2.
XX PD 08-JAN-2004.
XX PF 27-JUN-2003; 2003WO-US020347.
XX PR 28-JUN-2002; 2002US-0392431P.
XX PA (CENZ ) CENTOCOR INC.
XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX DR WPI; 2004-082870/08.
XX PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX PS Claim 2; SEQ ID NO 477; 129pp; English.
XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX SQ Sequence 11 AA;
XX Query Match 92.9%; Score 26; DB 8; Length 11;
XX Best Local Similarity 66.7%; Pred. No. 12;
XX Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB |||||
3 VRDQIWAAL 11

RESULT 11
ADJ51619
ID ADJ51619 standard; peptide; 11 AA.
XX AC ADJ51619;
XX DT 06-MAY-2004 (first entry)
XX XX
```

DE CH1 deleted mimetibody-related peptide SeqID477.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunological disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

OS Synthetic.

XX WO2004002424-A2.

PN 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

PI Kutoloshki KA;

XX WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for

CC diagnosing, preventing or treating cardiovascular, dermatologic,

CC endocrine, gastrointestinal, gynecologic, infectious, neurologic and

CC nutritional disorders.

XX Claim 14; SEQ ID NO 477; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CH1 deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,

CC obstetric, haematologic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX Sequence 11 AA;

XX Query Match 92.9%; Score 26; DB 8; Length 11;

XX Best Local Similarity 66.7%; Pred. No. 12;

XX Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXL 9

DB 3 VRDQIWAAL 11

Query Match 92.9%; Score 26; DB 8; Length 11;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 13
 AAY22407
 ID AAY22407 standard; peptide; 13 AA.
 XX
 AC AAY22407;
 XX
 XX 27-SEP-1999 (first entry)
 XX
 XX TPO receptor binding peptide sequence, SEQ ID NO. 58.
 XX
 XX TPO, thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
 XX
 XX Synthetic.
 XX
 XX US5932546-A.
 PN
 XX 03-AUG-1999.
 PD
 XX 04-OCT-1996; 96US-00726464.
 PF
 XX 04-OCT-1996; 96US-00726464.
 PR
 XX (GLAX) GLAXO WELLCOME INC.
 PA
 XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
 PI Cwirla SE, Johnson SS;
 PI
 XX WPI; 1999-457122/38.
 DR
 XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopenia.
 PT
 XX Disclosure; Col 15; 36pp; English.
 PS
 XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 CC mM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders or thrombocytopaenia,
 CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration
 CC
 XX Sequence 13 AA;
 SQ

Query Match 92.9%; Score 26; DB 2; Length 13;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIMLSL 11

RESULT 14
 AAY65258
 ID AAY65258 standard; protein; 95 AA.
 XX

Query Match 92.9%; Score 26; DB 3; Length 95;
 Best Local Similarity 55.6%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 |||||
 Db 66 VRDQLASSL 74

RESULT 15
 ADU72822
 ID ADU72822 standard; protein; 95 AA.
 XX
 AC ADU72822;
 XX
 DT 10-FEB-2005 (first entry)

AAV65258;
 01-FEB-2000 (first entry)
 Human 5' EST related polypeptide SEQ ID NO:1419.
 Human, 5' EST; expressed sequence tag; secreted protein; diagnosis;
 gene therapy; chromosome mapping; upstream regulatory sequence; forensic;
 location; development; protein synthesis; stability; regulation;
 identification.
 Homo sapiens.
 WO9953051-A2.
 21-OCT-1999.
 09-APR-1999; 99WO-1B000712.
 09-APR-1998; 98US-00057719.
 28-APR-1998; 98US-00069047.
 (GEST) GENSET.
 Dumas Milne Edwards J, Duclert A, Giordano J;
 WPI; 2000-038446/03.
 N-PSDB; AA242872.
 Novel secreted protein 5' expressed sequence tag sequences used in
 diagnostic, forensic, gene therapy, and chromosome mapping procedures.
 Claim 3; Page 779; 837pp; English.
 AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
 sequences, corresponding to human secreted proteins. AAY64651 to AAY65438
 represent the EST-related proteins corresponding to AA242265 to AA243052.
 The 5' ESTs can be used for producing secreted human gene products. They
 can be used to identify and isolate 5' untranslated regions (UTRs) and
 upstream regulatory regions which control the location, development
 stage, rate, and quantity of protein synthesis, as well as stability of
 mRNA. The ESTs are also useful as probes for chromosome mapping, and to
 obtain full length cDNA clones. The ESTs can also be used in forensic
 procedures to identify individuals, or in diagnostic procedures to
 identify individuals having genetic diseases resulting from abnormal gene
 expression. The products may also be used in gene therapy protocols. The
 nucleic acids encoding signal peptides can be used for directing
 extracellular secretion of a polypeptide or the insertion of a
 polypeptide into a membrane, or importing a polypeptide into a cell. The
 proteins encoded by the EST sequences may be useful in treating a variety
 of human conditions. Secreted proteins have therapeutic value, and the
 identification of new secreted proteins is valuable. AA242249 to AA242264
 and AAY64644 to AAY64650 represent sequences used in the exemplification
 of the present invention
 Sequence 95 AA;
 SQ

XX Signal peptide-containing polypeptide fragment, SEQ ID NO:1419.
XX Protein secretion; recombinant protein; diagnosis; mapping; forensic;
XX gene therapy; signal peptide.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..84
XX /label= Signal_peptide
XX Protein 85..95
XX /label= Mature_polypeptide_fragment
XX US6822072-B1.
XX 23-NOV-2004.
XX 21-DEC-1999; 99US-00471276.
XX 09-APR-1998; 98US-00057719.
XX 28-APR-1998; 98US-00069047.
XX 09-APR-1999; 99WO-IB000712.
XX (GEST) GENSET SA.
XX Edwards JDM, Duclert A, Giordano J;
XX WPI; 2004-812112/80.
XX N-PSDB; ADU72034.
XX New expressed sequence tags and encoded human proteins useful for
XX diagnosing, preventing or treating diseases such as autoimmune disorders,
XX inflammation, wounds or infections, or in forensic or chromosome mapping
XX procedures.
XX Example 15; SEQ ID NO 1419; 72pp; English.
XX The invention relates to an isolated or purified signal peptide
XX consisting of residues 1-16 of ADU7234 (signal peptide given separately
XX as ADU73026) which directs the extracellular secretion of a polypeptide
XX to which it is operably linked. The invention also relates to a method of
XX producing the signal peptide. The invention further discloses: isolated,
XX purified or enriched 5' expressed sequence tags (ESTs), many of which
XX encode all or a part of a secretory signal peptide; polypeptides encoded
XX by these ESTs (EST-related polypeptides); antibodies which recognize the
XX EST-related polypeptides; vectors and host cells comprising EST-related
XX nucleic acids of the invention; an array of ESTs; methods involving the
XX use of signal peptides of the invention to target polypeptides; and
XX methods involving the use of ESTs of the invention, for example, in
XX identifying a promoter in genomic DNA. The EST-encoded signal peptides of
XX the invention are useful for directing the secretion or import of a
XX recombinant polypeptide via the generation of protein fusions comprising
XX such signal peptides. The ESTs, EST-related polypeptides and methods of
XX the invention can be used for forensic procedures, chromosome mapping,
XX diagnostics, and therapeutic procedures, including gene therapy
XX Sequences ADU7215-ADU72919 represent incomplete polypeptides comprising
XX a signal peptide which are encoded by the 5' ESTs shown in ADU71427-
XX ADU72131. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from the US patent office at
XX seqdata.uspto.gov/sequence.html?DocID=US6822072.

XX Sequence 95 AA;
Query Match 92.9%; Score 26; DB 8; Length 95;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
Db 66 VRDQLASSL 74

RESULT 16
ADZ73813
ID ADZ73813 standard; protein; 95 AA.
XX
AC ADZ73813;
XX
DT 28-JUL-2005 (first entry)
XX
DE Human incomplete polypeptide including a signal peptide SEQ ID NO:1419.
XX expressed sequence tag; EST; expression; protein secretion; diagnostic;
XX forensic; gene therapy; haplotype mapping.
XX
OS Homo sapiens.
XX
PN US2005106595-A1.
XX
PD 19-MAY-2005.
XX
PF 25-AUG-2004; 2004US-00926683.
XX
PR 09-APR-1998; 98US-00057719.
XX
PR 28-APR-1998; 98US-00069047.
XX
PR 09-APR-1999; 99WO-IB000712.
XX
PR 21-DEC-1999; 99US-00471276.
XX
PA (GEST) GENSET SA.
XX
XX Dumas MEJ, Duclert A, Giordano J;
XX WPI; 2005-384300/39.
XX N-PSDB; ADZ73025.
XX
XX New purified nucleic acid expressing secreted proteins useful in
XX forensic, gene therapy, and chromosome mapping procedures, and diagnosing
XX or treating cancer, atherosclerosis and autoimmune diseases, diabetes,
XX asthma and infections.
XX
XX Claim 1; SEQ ID NO 1419; 79pp; English.
XX
XX The invention relates to a novel purified nucleic acid (I) comprising any
XX of (ADZ72418-ADZ73205) or (ADZ73994-ADZ74016) and their complements; at
XX least 15 consecutive nucleotides of (I) and their complements; or any of
XX 788 nucleotide sequences encoding fully defined sequences of 16-255 amino
XX acids (ADZ73206-ADZ73993). The invention discloses 5' EST's derived from
XX mRNAs encoding secreted proteins. The 5' EST's may be used to obtain
XX cDNAs and genomic DNAs corresponding to the 5' ESTs. The methods and
XX compositions of the present invention are useful for expressing secreted
XX proteins or its portions (claimed) or to obtain antibodies capable of
XX specifically binding to the secreted proteins, and in diagnostic,
XX forensic, gene therapy, and chromosome mapping procedures, and for
XX designing expression vectors and secretion vectors. The present sequence
XX is used in the exemplification of the invention. Note: The sequence data
XX for this patent is not represented in the printed specification but is
XX based on sequence information supplied in electronic format from the
XX USPTO web site seqdata.uspto.gov/sequence.html; Document ID: 20050106595.
XX
XX Sequence 95 AA;
Query Match 92.9%; Score 26; DB 9; Length 95;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
Db 66 VRDQLASSL 74

RESULT 17
ABU48350
ID ABU48350 standard; protein; 357 AA.
XX

AC ABU48350;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #33877.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Salmomella typhi.
OS
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX 06-SEP-2001; 2001US-00948993.
PR
XX 25-OCT-2001; 2001US-0342923P.
PR
XX 08-FEB-2002; 2002US-00072851.
PR
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR
XX N-PSDB; ACAS2220.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 76274; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 357 AA;
SQ

Query Match 92.9%; Score 26; DB 6; Length 357;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
Db 255 VRDQIAKYL 263

RESULT 18

ADI67225
ID ADI67225 standard; protein; 465 AA.

XX AC ADI67225;

XX DT 24-FEB-2005 (revised)

XX DT 22-APR-2004 (first entry)

XX DE Lactobacillus rhamnosus polypeptide sequence #99.

XX KW Lactic acid bacteria; vaccine; bacterial infection; microbe;

XX KW milk-derived product; food product; food additive;

XX KW nutritional supplement; bioactive substance; probiotic supplement;

XX KW flavour; aroma; texture; nutritional value; food; microorganism;

XX KW antibacterial.

XX OS Lactobacillus rhamnosus; strain HN001.

XX PN US2004009490-A1.

XX PD 15-JAN-2004.

XX PF 03-OCT-2002; 2002US-00264213.

XX PR 09-AUG-1999; 99US-0147852P.

XX PR 09-AUG-1999; 99US-0147853P.

XX PR 01-SEP-1999; 99US-0152031P.

XX PR 01-SEP-1999; 99US-0152032P.

XX PR 08-AUG-2000; 2000US-00634238.

XX PR 02-OCT-2001; 2001US-00971536.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Glenn M, Havukkala IJ, Lubbers M, Dekker J;

XX WPI; 2004-090459/09.

XX DR N-PSDB; ADI67100.

XX New polynucleotide from Lactobacillus rhamnosus HN001 strain, useful for
PT preparing a vaccine against bacterial infections or for modifying the
PT flavor, aroma or nutritional benefits of a bioactive or probiotic
PT supplement product.

XX PS Claim 21; SEQ ID NO 220; 54pp; English.

XX The present invention relates to the isolation of novel lactic acid
CC bacteria (Lactobacillus rhamnosus strain HN001) polynucleotide sequences,
CC and the proteins encoded by them. Also disclosed are oligonucleotide
CC probes and primers, and genetic constructs comprising the polynucleotide
CC sequences of the invention. The polynucleotide sequences are useful for
CC preparing a vaccine against bacterial infections or for improving the
CC properties of microbes used in the manufacture of milk-derived products,
CC food products, food additives, nutritional supplements, bioactive
CC substances or probiotic supplements, and for modifying the flavour,
CC aroma, texture and/or nutritional value of foods. They are also useful
CC for identifying microorganisms having a trait associated with the
CC polynucleotide. The present sequence represents a novel L. rhamnosus
CC polypeptide sequence of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification. The complete
CC sequence data for this patent was obtained in electronic format directly
CC from the USPTO web site at seqdata.uspto.gov.

XX SQ Sequence 465 AA;

Query Match 92.9%; Score 26; DB 8; Length 465;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;

DR WPI; 2005-028716/03.

XX New substantially purified *Myxococcus xanthus* nucleic acid molecule

PT encoding a nitrite reductase, useful for determining gene expression,

PT identifying mutations in a gene of interest, and for constructing

PT mutations in a gene of interest.

XX

XX Example 2; SEQ ID NO 16112; 25pp; English.

XX The invention relates to a substantially purified nucleic acid molecule

CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a

CC recombinant DNA construct for expression of a nitrite reductase gene in a

CC plant cell, and a plant cell comprising the recombinant DNA construct.

CC The nucleic acid is useful for determining gene expression, identifying

CC mutations in a gene of interest, and for constructing mutations in a gene

CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent

CC a group of 7134 *Myxococcus xanthus* proteins and peptides. Note: The

CC sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from USPTO

XX

XX Sequence 2559 AA;

Query Match 92.9%; Score 26; DB 9; Length 2559;

Best Local Similarity 66.7%; Pred. No. 5.5e+03;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 268 VRDQIRSVL 276

RESULT 22

AAAB16977

ID AAB16977 standard; peptide; 9 AA.

AC AAB16977;

DT 31-OCT-2000 (first entry)

DE TPO-mimetic peptide sequence SEQ ID NO:33.

KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;

KW autoimmune disease; cycostatic; antiasthmatic; thrombolytic; VEGF;

KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;

KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;

KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;

KW vascular endothelial growth factor; matrix metalloproteinase; asthma;

KW thrombosis; pharmaceutical.

OS Synthetic.

XX

XX WO200024782-A2.

XX

XX 04-MAY-2000.

XX

XX 25-OCT-1999; 99WO-US025044.

XX

XX 23-OCT-1998; 98US-0105371P.

XX

XX 22-OCT-1999; 99US-00428082.

XX

XX (AMGE-) AMGEN INC.

XX

XX Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX

XX Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.

XX

XX Claim 19; Page 207; 608pp; English.

XX

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC

CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each

CC independently selected from -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active

CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic,

CC thrombolytic and immunosuppressive activities. DNAs, vectors and host

CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,

CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAM69443 to AAM69526 and AAB16955 to

CC AAB18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention

XX

XX Sequence 9 AA;

Query Match 89.3%; Score 25; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 1 VRDQIXXXL 9

RESULT 23

ABB72863

ID ABB72863 standard; peptide; 9 AA.

XX

AC ABB72863;

XX

DT 05-APR-2002 (first entry)

XX

XX TPO mimetic peptide SEQ ID NO:33.

XX

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;

KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;

KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;

KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;

KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

KW sleep disorder; neurological degenerative disease; anaemia;

KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;

KW Fanconi's syndrome.

XX

OS Homo sapiens.

OS Synthetic.

XX

XX WO200183525-A2.

XX

XX 08-NOV-2001.

XX

XX 02-MAY-2001; 2001WO-US014310.

XX

XX 03-MAY-2000; 2000US-00563286.

XX

XX (AMGE-) AMGEN INC.

XX

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX

XX Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention
 XX

Sequence 9 AA;

Query Match 89.3%; Score 25; DB 5; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 |||||
 DB 1 VRDQIXXXL 9

RESULT 24

ABG71745
 ID ABG71745 standard; peptide; 9 AA.

AC ABG71745;

XX 20-JAN-2003 (first entry)

XX TPO receptor, MPL, agonist peptide consensus sequence #1.

XX Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
 KW complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
 KW thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
 KW T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
 KW proliferation; growth; differentiation; haematopoietic cell;
 KW platelet progenitor cell; immune disorder; thrombocytopenia;
 KW disseminated intravascular coagulation; stem cell; transplantation;
 KW gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 6..8 /label= Unknown
 FT /notes "Residues are not defined within the
 FT specification"

XX WO200278612-A2.

XX 10-OCT-2002.

XX 02-APR-2002; 2002WO-US010301.

XX 02-APR-2001; 2001US-0281183P.

XX (PURD) PURDUE PHARMA LP.

XX Soltis DA, Burch RM, Ogert RA;

XX WPI; 2003-040615/03.

XX New thrombopoietin synthebodies, useful for stimulating proliferation,
 PT growth, or differentiation of hematopoietic cells, for treating or
 PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.

XX Disclosure; Page 11; 97pp; English.

XX The invention discloses a variant of an immunoglobulin (Ig) variable
 CC heavy or light chain domain that comprises at least one complementarity
 CC determining region (CDR) and framework regions flanking the CDR. The CDR
 CC also has added or substituted to it, at least one binding sequence which
 CC is heterologous to the CDR and is an antigenic, agonistic sequence from a
 CC thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
 CC can be a binding sequence heterologous to the CDR, a cytotoxic T-
 CC lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
 CC cell sequence or a combination of each. The variant or thrombopoietin
 CC synthebody, pharmaceutical and vaccine compositions are useful for
 CC stimulating proliferation, growth or differentiation of haematopoietic
 CC cells, particularly platelet progenitor cells. The variants are also
 CC useful for treating or preventing haematopoietic or immune disorders
 CC resulting from chemotherapy, radiation therapy, or bone marrow
 CC transfusions (e.g. thrombocytopenia or disseminated intravascular
 CC coagulation). Compositions comprising the synthebodies can be used for
 CC the mobilisation, amplification and ex vivo expansion of stem cells and
 CC committed precursor cells for autologous and allogeneic transplantation
 CC as well as for the expansion of stem cells for gene therapy. They are
 CC also useful as diagnostic or analytical reagents for studying the
 CC function of thrombopoietin and its receptor in vivo or in vitro. The
 CC sequence presented is the TPO receptor (MPL) agonist peptide consensus
 CC sequence #1

XX Sequence 9 AA;

Query Match 89.3%; Score 25; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 |||||
 DB 1 VRDQIXXXL 9

RESULT 25

ADJ73014
 ID ADJ73014 standard; peptide; 9 AA.

XX ADJ73014;

XX 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 468.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW TPO.

XX Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallan BJ, Ghayeb J;

DR WPI; 2003-804237/75.
 XX
 PT New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX
 XX Disclosure; SEQ ID NO 468; 97pp; English.
 XX
 CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is a TPO mimetic peptide sequence used to make a
 CC mimetibody of the invention.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 89.3%; Score 25; DB 7; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VRDQIXXXL 9
 Db 1 VRDQIXXXL 9
 |||||
 |||||
 RESULT 26
 ADJ52649
 ID ADJ52649 standard; peptide; 9 AA.
 XX
 AC ADJ52649;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE CH1 deleted mimetibody-related peptide SeqID468.
 XX
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nontropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..9
 FT /note= "All Xaa's in this sequence are unidentified amino
 FT acids"
 FT
 XX WO2004002417-A2.
 PN
 XX
 XX 08-JAN-2004.
 PD
 XX
 XX 27-JUN-2003; 2003WO-US020347.
 PF
 XX 28-JUN-2002; 2002US-0392431P.
 PR
 XX (CENZ) CENTOCOR INC.
 PA
 XX

PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutoloski KA;
 XX
 XX WPI; 2004-082870/08.
 DR
 XX
 XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 PT
 XX Claim 2; SEQ ID NO 468; 129pp; English.
 PS
 XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nontropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.
 XX
 XX Sequence 9 AA;
 SQ
 Query Match 89.3%; Score 25; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VRDQIXXXL 9
 Db 1 VRDQIXXXL 9
 |||||
 |||||
 RESULT 27
 ADJ51610
 ID ADJ51610 standard; peptide; 9 AA.
 XX
 AC ADJ51610;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE CH1 deleted mimetibody-related peptide SeqID468.
 XX
 KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunologic disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 1..9
 FT /note= "All Xaa's in this sequence are unidentified amino
 FT acids"
 FT
 XX WO2004002424-A2.
 PN

XX 08-JAN-2004.
 XX 30-JUN-2003; 2003WO-US020495.
 XX 28-JUN-2002; 2002US-0392431P.
 XX 19-SEP-2002; 2002US-0412144P.
 XX (CENZ) CENTOCOR INC.
 XX Heavner GA, Knight DM, Ghayeb J, Scallon BJ, Nesspor TC;
 XX Kutoloski KA;
 XX WPI; 2004-082872/08.
 XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 XX diagnosing, preventing or treating cardiovascular, dermatologic,
 XX endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 XX nutritional disorders.
 XX Claim 14; SEQ ID NO 468; 123pp; English.
 XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 XX which encode them), compositions, methods and uses. The invention may be
 XX useful for the development of compounds with an osteopathic,
 XX cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 XX gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 XX immunomodulator, antiallergic, muscular-Gen, cyrostatic,
 XX antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
 XX respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 XX modulator or cytokine-agonist. The methods and compositions of the
 XX present invention are useful for the diagnosis, prevention and/or
 XX treatment of diseases or conditions associated with aberrant expression
 XX or activity of the CHI deleted mimetibody, such as a bone or joint,
 XX cardiovascular, dental or oral, dermatological, ear, nose or throat,
 XX endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 XX obstetric, haematologic, immunological, allergic, infectious,
 XX musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 XX pediatric, psychiatric, renal or pulmonary disorders. The present
 XX sequence is that of a peptide which may be used during the creation of a
 XX mimetibody of the invention.
 XX Sequence 9 AA;
 Query Match 89.3%; Score 25; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRDQIXXXL 9
 Db 1 VRDQIXXXL 9
 RESULT 28
 AAY22357
 ID AAY22357 standard; peptide; 10 AA.
 XX AC AAY22357;
 XX 27-SEP-1999 (first entry)
 XX TPO receptor binding peptide sequence, SEQ ID NO. 8.
 XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 XX haematological disorder; therapy; bone marrow transfusion; diagnosis;
 XX haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
 XX Synthetic.
 XX US5932546-A.
 XX 03-AUG-1999.
 XX New low molecular weight thrombopoietin agonists, particularly peptides,
 XX for treatment of hematological disease and thrombocytopenia.

PF 04-OCT-1996; 96US-00726464.
 XX 04-OCT-1996; 96US-00726464.
 XX (GLAX) GLAXO WELLCOME INC.
 XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
 XX Cwirla SE, Johnson SS;
 XX WPI; 1999-457122/38.
 XX New low molecular weight thrombopoietin agonists, particularly peptides,
 XX for treatment of hematological disease and thrombocytopenia.
 XX Disclosure; Col 13-14; 36pp; English.
 XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 XX peptide of the invention. The peptide has: (i) a molecular weight below
 XX 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 XX mM. The peptides are used to treat conditions requiring a thrombopoietin
 XX agonist, particularly haematological disorders or thrombocytopaenia,
 XX especially resulting from chemotherapy, radiation therapy or bone marrow
 XX transfusions. Also when labelled they may be used for diagnosis
 XX (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 XX in vitro expansion of megakaryocytes and committed progenitor cells, and
 XX for the development/identification of other TR agonists. The compounds
 XX accelerate thrombocyte regeneration
 XX Sequence 10 AA;
 Query Match 89.3%; Score 25; DB 2; Length 10;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VRDQIXXXL 9
 Db 2 VRDQVSWAL 10
 RESULT 29
 AAY22360
 ID AAY22360 standard; peptide; 10 AA.
 XX AC AAY22360;
 XX 27-SEP-1999 (first entry)
 XX TPO receptor binding peptide sequence, SEQ ID NO. 11.
 XX TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 XX haematological disorder; therapy; bone marrow transfusion; diagnosis;
 XX haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
 XX Synthetic.
 XX US5932546-A.
 XX 03-AUG-1999.
 XX 04-OCT-1996; 96US-00726464.
 XX 04-OCT-1996; 96US-00726464.
 XX (GLAX) GLAXO WELLCOME INC.
 XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
 XX Cwirla SE, Johnson SS;
 XX WPI; 1999-457122/38.
 XX New low molecular weight thrombopoietin agonists, particularly peptides,
 XX for treatment of hematological disease and thrombocytopenia.

PS Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding peptide of the invention. The peptide has: (i) a molecular weight below 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10 mM. The peptides are used to treat conditions requiring a thrombopoietin agonist, particularly haematological disorders or thrombocytopaenia, especially resulting from chemotherapy, radiation therapy or bone marrow transplantations. Also when labelled they may be used for diagnosis (detecting TR on cells); for studying mechanisms of haematopoiesis; for in vitro expansion of megakaryocytes and committed progenitor cells, and for the development/identification of other TR agonists. The compounds accelerate thrombocyte regeneration

XX SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 2; Length 10;

Best Local Similarity 55.6%; Pred. No. 18;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
||||: |

Db 2 VRDQVSWAL 10

RESULT 30

ID AAB16981 standard; peptide; 10 AA.

XX AAB16981;

XX 31-OCT-2000 (first entry)

XX TPO-mimetic peptide sequence SEQ ID NO:37.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer; autoimmune disease; cycostatic; antiasthmatic; thrombolytic; VEGF; immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP inhibitor; erythropoietin; thrombopoietin; interleukin 1; cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor; vascular endothelial growth factor; matrix metalloproteinase; asthma; thrombosis; pharmaceutical.

XX Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC, Gudas JM;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 208; 608pp; English.

XX The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 3; Length 10;

Best Local Similarity 55.6%; Pred. No. 18;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
||||: |

Db 2 VRDQVSWAL 10

RESULT 31

ABB72867

ID ABB72867 standard; peptide; 10 AA.

XX ABB72867;

XX 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:37.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP; TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianemic; anorectic; antiinfertility; haemostatic; dermatological; neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cycostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (II), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 89.3%; Score 25; DB 5; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 2 VRDQVSWAL 10
||||: |

RESULT 32
ADJ73018
ID ADJ73018 standard; peptide; 10 AA.

XX
AC ADJ73018;

DT 06-MAY-2004 (first entry)

DE TPO mimetic peptide sequence SeqID 472.

KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 472; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurological
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 7; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 2 VRDQVSWAL 10
||||: |

RESULT 33
ADJ52653

ID ADJ52653 standard; peptide; 10 AA.

XX
AC ADJ52653;

XX 06-MAY-2004 (first entry)

DE CH1 deleted mimetibody-related peptide SeqID472.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

XX WO2004002417-A2.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US020347.

XX 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;

XX WPI; 2004-082870/08.

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 472; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,

CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 8; Length 10;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
 ||||: |
 Db 2 VRDQVSWAL 10

RESULT 34
 ADJ51614
 ID ADJ51614 standard; peptide; 10 AA.

XX ADJ51614;

DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID472.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

XX WO2004002424-A2.

XX 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

XX 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
 PI Kurolooski KA;

XX WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.

XX Claim 14; SEQ ID NO 472; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX Sequence 10 AA;

Query Match 89.3%; Score 25; DB 8; Length 10;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
 ||||: |
 Db 2 VRDQVSWAL 10

RESULT 35

AAB16989
 ID AAB16989 standard; peptide; 11 AA.

XX AAB16989;

DT 31-OCT-2000 (first entry)

XX TPO-mimetic peptide sequence SEQ ID NO.45.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.

XX Synthetic.

XX WO200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 211; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 11 AA;

Query Match 89.3%; Score 25; DB 3; Length 11;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIXXXL 11

RESULT 36

ABB72875
 ID ABB72875 standard; peptide; 11 AA.

XX AC ABB72875;

XX DT 05-APR-2002 (first entry)

XX DE TPO mimetic peptide SEQ ID NO:45.

XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytosolic; antirheumatic; antiarthritic; haemostatic; ophthalmological;
 KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200183525-A2.

XX PD 08-NOV-2001.

XX PF 02-MAY-2001; 2001WO-US014310.

XX PR 03-MAY-2000; 2000US-00563286.

XX PA (AMGE-) AMGEN INC.

XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX DR WPI; 2002-130313/17.

XX PT Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX PS Claim 39; Page 43; 176pp; English.

XX CC The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 11 AA;

Query Match 89.3%; Score 25; DB 5; Length 11;

Best Local Similarity 100.0%; Pred. No. 21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 3 VRDQIXXXL 11

RESULT 37

ADJ52661
 ID ADJ52661 standard; peptide; 11 AA.

XX AC ADJ52661;

XX DT 06-MAY-2004 (first entry)

XX DE CH1 deleted mimetibody-related peptide SeqID480.

XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.

XX OS Unidentified.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1..9
 FT /note= "All Xaa's in this sequence are unidentified amino
 FT acids"

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX PA Heavner GA, Knight DM, Ghayeb J, Scallon BJ, Nesspor TC;

XX PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.

XX Claim 2; SEQ ID NO 480; 139pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX Sequence 11 AA;

Query Match 89.3%; Score 25; DB 8; Length 11;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 Db 3 VRDQIXXXL 11
 |||||

RESULT 38

AAAY22388
 ID AAY22388 standard; peptide; 15 AA.

XX AC AAY22388;

XX DT 27-SEP-1999 (first entry)

XX DE TPO receptor binding peptide sequence, SEQ ID NO. 39.

XX KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX OS Synthetic.

XX PN US5932546-A.

XX PD 03-AUG-1999.

XX PF 04-OCT-1996; 96US-00726464.

XX PR 04-OCT-1996; 96US-00726464.

XX PA (GLAX) GLAXO WELLCOME INC.

PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

PI Cwirla SE, Johnson SS;

XX WPI; 1999-457122/38.

XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopaenia.

XX Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10

CC mM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders or thrombocytopaenia,
 CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration

XX Sequence 15 AA;

Query Match 89.3%; Score 25; DB 2; Length 15;
 Best Local Similarity 55.6%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 Db 6 VRDQVLAFL 14
 |||||

RESULT 39

AAAY22375
 ID AAY22375 standard; peptide; 15 AA.

XX AC AAY22375;

XX DT 27-SEP-1999 (first entry)

XX DE TPO receptor binding peptide sequence, SEQ ID NO. 26.

XX KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX OS Synthetic.

XX PN US5932546-A.

XX PD 03-AUG-1999.

XX PF 04-OCT-1996; 96US-00726464.

XX PR 04-OCT-1996; 96US-00726464.

XX PA (GLAX) GLAXO WELLCOME INC.

PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

PI Cwirla SE, Johnson SS;

XX WPI; 1999-457122/38.

XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopaenia.

XX Disclosure; Col 13-14; 36pp; English.

XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 CC mM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders or thrombocytopaenia,
 CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration

XX Sequence 15 AA;

Query Match 89.3%; Score 25; DB 2; Length 15;
 Best Local Similarity 55.6%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX XX WPI; 2003-615309/58.
XX DR N-PSDB; ABD01529.
XX XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX XX Disclosure; SEQ ID NO 16704; 455pp; English.
XX CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-*P. aeruginosa* drugs, as templates for recombinant
XX CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
XX CC infection, and in detection of *P. aeruginosa* sequences or other sequences
XX CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 219 AA;
Query Match 89.3%; Score 25; DB 7; Length 219;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 1 VRDQIXXXL 9
Db 113 VRDQVRGAL 121
RESULT 43
AB061401
ID ABB61401 standard; protein; 250 AA.
AC ABB61401;
XX XX 26-MAR-2002 (first entry)
XX DE *Drosophila melanogaster* polypeptide SEQ ID NO 10995.
XX KW *Drosophila*; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX XX *Drosophila melanogaster*.
XX OS WO200171042-A2.
XX PN 27-SEP-2001.
XX PD 23-MAR-2001; 2001WO-US009231.
XX PF 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX XX (PEKE) PE CORP NY.
XX PA Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR N-PSDB; ABL05504.
XX XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX XX Disclosure; SEQ ID NO 10995; 21pp + Sequence Listing; English.
XX PS The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (AB057737-
XX CC AB072072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 250 AA;
Query Match 89.3%; Score 25; DB 4; Length 250;
Best Local Similarity 55.6%; Pred. No. 7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 1 VRDQIXXXL 9
Db 220 IRDQIAERL 228
RESULT 44
AAY43936
ID AAY43936 standard; protein; 276 AA.
XX XX AAY43936;
XX AC AAY43936;
XX DT 21-DEC-1999 (first entry)
XX DE Human protein kinase #6.
XX XX Prediction; secondary structure; alignment; evolutionary conservation;
XX KW homology; periodicity; co-variation analysis; antigenic site;
XX KW site directed mutagenesis; interaction.
XX OS Homo sapiens.
XX XX US5958784-A.
XX PN 28-SEP-1999.
XX PD 25-MAR-1992; 92US-00857224.
XX PF 25-MAR-1992; 92US-00857224.
XX PR (BENN/) BENNER S A.
XX PA Benner SA;
XX PI WPI; 1999-570766/48.
XX DR Predicting the folded structure of proteins.
XX XX Disclosure; Col 209-210; 113pp; English.
XX XX Sequences AAY43902-Y44015 represent proteins used in a novel method of
XX CC predicting the folded structure of proteins, by aligning sequences of
XX CC homologous proteins and using patterns of evolutionarily conserved and
XX CC varied sequences to assign positions. Positions in the alignment are
XX CC assigned to the surface or inside of the folded structure, active sites,
XX CC and parsing segments. Secondary structural units are assigned by
XX CC identifying periodicity in the assignments, and assembled into globular
XX CC form using distance constraints imposed by disulfide bridges, active site
XX CC assignments and co-variation analysis. The predicted secondary structures
XX CC are useful for identifying antigenic sites on a protein molecule, as
XX CC guides for site directed mutagenesis studies, and for understanding the

CC interaction of a protein with other molecules
XX
SQ Sequence 276 AA;

Query Match 89.3%; Score 25; DB 2; Length 276;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | : |
Db 82 VRDQVI AVL 90

RESULT 45
ADI57223
ID ADI57223 standard; protein; 314 AA.
XX
AC ADI57223;
XX
DT 22-APR-2004 (first entry)
XX
DE CDC7L1 dominant negative mutant amino acid sequence CDC7Sc.
XX
KW cell cycle arrest; CK2-specific siRNA molecule; short interfering RNA;
KW CK2 inhibition; PIM1-specific siRNA molecule; PIM1 inhibition;
KW Hb01-specific siRNA molecule; Hb01 inhibition; cytostatic; gene therapy;
KW cancer; proliferative disorder; human; CDC7L1; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
FN WO2004007754-A2.
XX
PD 22-JAN-2004.
XX
PF 14-JUL-2003; 2003WO-US022164.
XX
PR 12-JUL-2002; 2002US-0395443P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Hitoshi Y, Jenkins Y, Markovtsov V;
XX
PS WPI; 2004-122975/12.
XX

Identifying a compound that modulates cell cycle arrest, useful for developing therapeutic reagents for treating cancer comprising contacting a cell comprising a target polypeptide with the compound.

Disclosure; Fig 23; 180pp; English.

The present invention describes a method for identifying a compound (C) that modulates cell cycle arrest. The method comprises contacting a cell comprising a target polypeptide with the compound (C), where the target polypeptide encoded by the complement of a nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a polypeptide having an amino acid sequence selected from 18 148-1408 amino acid sequences (SEQ ID NO: 2-36, even numbers only). Also described: (1) modulating cell cycle arrest in a subject; (2) a CK2-specific short interfering RNA (siRNA) molecule comprising the sequence: (I) AACATGGAATTAGATCCACGT, where the siRNA molecule is from 21-30 nucleotide base pairs in length; (3) inhibiting expression of a CK2 gene in a cell; (4) a PIM1-specific siRNA molecule comprising the sequence: (II) AAAACTCCGAGTGAAGTGC, where the siRNA molecule is from 21-30 nucleotide base pairs in length; (5) inhibiting expression of a PIM1 gene in a cell; (6) an Hb01-specific siRNA molecule comprising the sequence: (III) AACTGAGCAAGTGGTGGATT, where the siRNA molecule is from 21-30 nucleotide base pairs in length; and (7) inhibiting expression of an Hb01 gene in a cell. (C) has cytostatic activity, and can be used in gene therapy. The method is useful for identifying a compound (c) that modulates cell cycle arrest. Compounds that modulate cell cycle arrest are useful for developing therapeutic reagents for treating cancer and other proliferative disorders. The present sequence represents a dominant

CC negative mutant of human CDC7L1, which is used in the exemplification of the present invention.
XX
SQ Sequence 314 AA;

Query Match 89.3%; Score 25; DB 8; Length 314;
Best Local Similarity 55.6%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | : |
Db 112 VRDQVI AVL 120

RESULT 46
ADI57229
ID ADI57229 standard; protein; 314 AA.
XX
AC ADI57229;
XX
DT 22-APR-2004 (first entry)
XX
DE CDC7L1 dominant negative mutant amino acid sequence.
XX
KW cell cycle arrest; CK2-specific siRNA molecule; short interfering RNA;
KW CK2 inhibition; PIM1-specific siRNA molecule; PIM1 inhibition;
KW Hb01-specific siRNA molecule; Hb01 inhibition; cytostatic; gene therapy;
KW cancer; proliferative disorder; human; CDC7L1; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
FN WO2004007754-A2.
XX
PD 22-JAN-2004.
XX
PF 14-JUL-2003; 2003WO-US022164.
XX
PR 12-JUL-2002; 2002US-0395443P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Hitoshi Y, Jenkins Y, Markovtsov V;
XX
PS WPI; 2004-122975/12.
XX

Identifying a compound that modulates cell cycle arrest, useful for developing therapeutic reagents for treating cancer comprising contacting a cell comprising a target polypeptide with the compound.

Example 18; Fig 26; 180pp; English.

The present invention describes a method for identifying a compound (C) that modulates cell cycle arrest. The method comprises contacting a cell comprising a target polypeptide with the compound (C), where the target polypeptide encoded by the complement of a nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a polypeptide having an amino acid sequence selected from 18 148-1408 amino acid sequences (SEQ ID NO: 2-36, even numbers only). Also described: (1) modulating cell cycle arrest in a subject; (2) a CK2-specific short interfering RNA (siRNA) molecule comprising the sequence: (I) AACATGGAATTAGATCCACGT, where the siRNA molecule is from 21-30 nucleotide base pairs in length; (3) inhibiting expression of a CK2 gene in a cell; (4) a PIM1-specific siRNA molecule comprising the sequence: (II) AAAACTCCGAGTGAAGTGC, where the siRNA molecule is from 21-30 nucleotide base pairs in length; (5) inhibiting expression of a PIM1 gene in a cell; (6) an Hb01-specific siRNA molecule comprising the sequence: (III) AACTGAGCAAGTGGTGGATT, where the siRNA molecule is from 21-30 nucleotide base pairs in length; and (7) inhibiting expression of an Hb01 gene in a cell. (C) has cytostatic activity, and can be used in gene therapy. The method is useful for identifying a compound (c) that modulates cell cycle arrest. Compounds that modulate cell cycle arrest are useful for developing therapeutic reagents for treating cancer and other

CC proliferative disorders. The present sequence represents a dominant
CC negative mutant of human CDC/11, which is used in the exemplification of
CC the present invention.

XX SQ Sequence 314 AA;
Query Match 89.3%; Score 25; DB 8; Length 314;
Best Local Similarity 55.6%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 112 VRDQVI AVL 120

RESULT 47
ADN18947
ID ADN18947 standard; protein; 507 AA.

XX AC ADN18947;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #1600.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WIPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 1600; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 507 AA;

Query Match 89.3%; Score 25; DB 8; Length 507;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 112 VRDQVI AVL 120

RESULT 48

ADN42509
ID ADS42509 standard; protein; 519 AA.

XX AC ADS42509;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #20939.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WIPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 20939; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 519 AA;

Query Match 89.3%; Score 25; DB 8; Length 519;
Best Local Similarity 55.6%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
|||||
DB 82 VRDQIATEI 90

RESULT 49
ABU22938
ID ABU22938 standard; protein; 530 AA.

XX AC ABU22938;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #8465.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Bordetella pertussis.

XX FN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA26808.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 50862; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid.

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway; (8)
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 530 AA;

Query Match 89.3%; Score 25; DB 6; Length 530;

Best Local Similarity 55.6%; Pred. No. 1.6e+03;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

DB 41 VRDQVAAQL 49

RESULT 50

ABO76610
ID ABO76610 standard; protein; 564 AA.

XX AC ABO76610;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #8785.

XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX PN US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX DR WPI; 2003-615309/58.

XX DR N-PSDB; ABD10181.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 25356; 455pp; English.
 XX
 CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 564 AA;

Query Match 89.3%; Score 25; DB 7; Length 564;
 Best Local Similarity 55.6%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 VRDQIXXL 9
 Db 234 VRDQVADL 242
 |||||
 |

Search completed: May 12, 2006, 10:37:49
 Job time : 94.4846 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 10.9231 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	543	2 T02022	hypothetical prote
2	27	96.4	1113	2 T47381	hypothetical prote
3	27	96.4	1198	2 E86402	hypothetical prote
4	27	96.4	1201	2 F86386	hypothetical prote
5	27	96.4	1305	2 H96559	hypothetical prote
6	27	96.4	1472	2 A84470	hypothetical prote
7	27	96.4	1756	2 T02599	hypothetical prote
8	26	92.9	211	2 T01194	hypothetical prote
9	26	92.9	357	2 AD1062	protein kinase [im
10	26	92.9	1236	2 B36329	hypothetical prote
11	26	92.9	1312	2 D85066	hypothetical prote
12	25	89.3	115	2 S09866	hypothetical prote
13	25	89.3	193	2 D83620	hypothetical prote
14	25	89.3	243	2 T23381	hypothetical prote
15	25	89.3	337	2 S14268	peroxidase (EC 1.1
16	25	89.3	507	2 A25228	protein kinase CDC
17	25	89.3	610	2 F82192	ABC transporter, A
18	25	89.3	621	2 AF3016	Na+/H+ antiporter
19	25	89.3	642	2 B98268	probable sodium/hy
20	25	89.3	686	2 F89895	ATP-dependent DNA
21	25	89.3	780	2 T00122	hypothetical prote
22	25	89.3	1430	2 T21910	hypothetical prote
23	25	89.3	3871	2 T22812	hypothetical prote
24	24	85.7	103	2 G82442	Sufl family protei
25	24	85.7	109	2 A87425	hypothetical prote
26	24	85.7	166	2 T39825	hypothetical prote
27	24	85.7	177	2 B36662	hypothetical prote
28	24	85.7	185	2 S66113	hypothetical prote
29	24	85.7	193	1 S39401	hydrogenase matura

transcription regu	205	85.7	24	30
type iii secretion	235	85.7	24	31
1-acylglycerol-3-p	250	85.7	24	32
Conserved hypothet	257	85.7	24	33
ketoacyl reductase	259	85.7	24	34
probable RNA pseud	260	85.7	24	35
hypothetical prote	260	85.7	24	36
hypothetical prote	260	85.7	24	37
hypothetical prote	260	85.7	24	38
carboxypeptidase B	302	85.7	24	39
conserved hypothet	303	85.7	24	40
general secretion	315	85.7	24	41
hypothetical prote	324	85.7	24	42
hypothetical prote	351	85.7	24	43
hypothetical prote	369	85.7	24	44
hypothetical prote	376	85.7	24	45
hypothetical prote	448	85.7	24	46
hypothetical prote	471	85.7	24	47
hypothetical prote	505	85.7	24	48
ferredoxin-nitrite	512	85.7	24	49
hypothetical prote	568	85.7	24	50
excinnuclease ABC c	600	85.7	24	51
medium-chain acyl-	600	85.7	24	52
cell division prot	628	85.7	24	53
pepidylprolyl iso	628	85.7	24	54
probable oxidoredu	648	85.7	24	55
dnak-type molecula	654	85.7	24	56
methyl-accepting c	781	85.7	24	57
hypothetical prote	788	85.7	24	58
methyl-accepting c	891	85.7	24	59
hypothetical prote	997	85.7	24	60
p63 related protei	1114	85.7	24	61
hypothetical prote	1175	85.7	24	62
hypothetical prote	1175	85.7	24	63
hypothetical prote	1203	85.7	24	64
probable retroviri	1489	85.7	24	65
hypothetical prote	1633	85.7	24	66
gas vesicle protein	83	82.1	23	67
heat shock protein	91	82.1	23	68
protein F28K20.11	140	82.1	23	69
conserved hypothet	149	82.1	23	70
hypothetical prote	154	82.1	23	71
CT676 hypothetical	171	82.1	23	72
conserved hypothet	171	82.1	23	73
hypothetical prote	173	82.1	23	74
conserved hypothet	177	82.1	23	75
hypothetical prote	180	82.1	23	76
probable hydantoin	203	82.1	23	77
hypothetical prote	236	82.1	23	78
hypothetical prote	248	82.1	23	79
hypothetical prote	265	82.1	23	80
riboflavin kinase/	303	82.1	23	81
probable sn2-acylg	303	82.1	23	82
ABC transporter, A	308	82.1	23	83
tetratricopeptide	317	82.1	23	84
5-methyltetrahydro	319	82.1	23	85
UDP-glucose 4-epim	328	82.1	23	86
conserved hypothet	331	82.1	23	87
D-lactate dehydrog	333	82.1	23	88
transcription regu	334	82.1	23	89
D-lactate dehydrog	337	82.1	23	90
hypothetical prote	349	82.1	23	91
phage-related prot	360	82.1	23	92
exo-alpha-bialidas	382	82.1	23	93
matng factor MAR1	383	82.1	23	94
bacteriophage N4 a	389	82.1	23	95
UDP-N-acetyl gluco	390	82.1	23	96
UDP-N-acetyl gluco	390	82.1	23	97
sensor histidine k	392	82.1	23	98
tocopherol associa	405	82.1	23	99
transcription regu	405	82.1	23	100
N-carbamyl-L-amino	409	82.1	23	101
probable Tail-like	409	82.1	23	102

103	23	82.1	426	2	H71636	3-oxoacyl-(acyl-ca	176	22	78.6	137	2	T20957	hypothetical prote
104	23	82.1	426	2	D87457	3-oxoacyl-(acyl-ca	177	22	78.6	139	2	C69950	arsenate reductase
105	23	82.1	427	2	A84155	hypothetical prote	178	22	78.6	139	2	T33713	conserved hypoteth
106	23	82.1	428	2	B97848	hypothetical prote	179	22	78.6	147	2	I46623	rearranged r-cell
107	23	82.1	433	2	A84619	probable serine ca	180	22	78.6	151	2	Ar0014	conserved hypoteth
108	23	82.1	446	2	F82609	aminopeptidase P X	181	22	78.6	154	2	E82945	galactoside O-acet
109	23	82.1	451	2	F87659	conserved hypoteth	182	22	78.6	160	2	G95113	regulatory protein
110	23	82.1	460	2	I57546	Rabin3 - rat	183	22	78.6	160	2	H97982	protein F7F22.12 l
111	23	82.1	464	2	T24011	hypothetical prote	184	22	78.6	169	2	C96494	conserved hypoteth
112	23	82.1	470	2	C70641	hypothetical prote	185	22	78.6	196	2	D82143	hypothetical prote
113	23	82.1	495	2	T52515	related to n-alkan	186	22	78.6	201	2	AC2371	hypothetical prote
114	23	82.1	503	1	A35376	cholesterol 7alpha	187	22	78.6	204	1	MNNZHS	nonstructural prot
115	23	82.1	504	2	S39399	hypothetical prote	188	22	78.6	204	1	MNNZSV	hypothetical prote
116	23	82.1	504	2	G84251	hypothetical prote	189	22	78.6	208	2	A82077	MutT/nudix family
117	23	82.1	527	2	D70595	probable ATP-depen	190	22	78.6	209	2	A71313	probable V-type AT
118	23	82.1	536	2	E82559	virulence factor X	191	22	78.6	212	2	G82176	probable outer mem
119	23	82.1	544	2	F87010	probable ATP-depen	192	22	78.6	217	2	T37039	hypothetical prote
120	23	82.1	557	2	E89901	conserved hypoteth	193	22	78.6	217	2	S66155	probable DNA-direc
121	23	82.1	566	2	B28391	alpha-amylase (EC	194	22	78.6	221	2	D87469	transcription regu
122	23	82.1	566	2	S14063	alpha-amylase (EC	195	22	78.6	222	2	AD1397	ABC transporter, A
123	23	82.1	566	2	C70482	repair excision nu	196	22	78.6	223	2	AG1772	ABC transporter, A
124	23	82.1	575	1	VCLJHD	env polyprotein pr	197	22	78.6	232	2	C95245	hypothetical prote
125	23	82.1	575	2	S44018	aerobactin biosynt	198	22	78.6	232	2	H98109	2,3,4,5-tetrahydro
126	23	82.1	579	2	F69157	excinuclease ABC C	199	22	78.6	232	2	A69131	conserved hypoteth
127	23	82.1	588	1	BVEUC	excinuclease ABC,	200	22	78.6	236	2	A11558	tetrahydrodipicoli
128	23	82.1	588	2	D85808	excinuclease ABC,	201	22	78.6	236	2	F69866	tetrahydrodipicoli
129	23	82.1	588	2	C90960	excinuclease ABC s	202	22	78.6	236	2	AC1201	tetrahydrodipicoli
130	23	82.1	592	2	D95393	hypothetical prote	203	22	78.6	236	2	H71056	probable arylmalon
131	23	82.1	593	2	T44479	conserved hypoteth	204	22	78.6	239	2	B83296	probable transcrip
132	23	82.1	594	2	S09857	hypothetical prote	205	22	78.6	239	2	T34733	hypothetical prote
133	23	82.1	598	2	AG0890	probable arylsulfa	206	22	78.6	240	2	E83983	tetrahydrodipicoli
134	23	82.1	610	2	AD0749	excinuclease ABC C	207	22	78.6	245	2	S43565	R01H10.4 protein (
135	23	82.1	610	2	AG0277	excinuclease ABC C	208	22	78.6	252	2	E86950	conserved hypoteth
136	23	82.1	610	2	S77337	ABC-type transport	209	22	78.6	256	2	A86660	acetyltransferase
137	23	82.1	611	2	T44880	probable acyl-coA	210	22	78.6	261	2	B83969	RNA polymerase spo
138	23	82.1	611	2	H70938	probable fadB5 pro	211	22	78.6	267	2	S15486	probable ABC-type
139	23	82.1	613	1	S74461	ABC transporter sl	212	22	78.6	269	2	B71365	probable phosphome
140	23	82.1	620	2	A83182	hypothetical prote	213	22	78.6	270	2	D90113	hypothetical prote
141	23	82.1	626	1	S77286	excinuclease ABC C	214	22	78.6	276	2	F87679	chemotaxis protein
142	23	82.1	627	2	A12146	excinuclease ABC C	215	22	78.6	278	2	AD0146	probable ABC trans
143	23	82.1	669	2	D82739	excinuclease ABC s	216	22	78.6	280	2	T09576	Lbx1 transcription
144	23	82.1	674	2	T42071	hypothetical prote	217	22	78.6	281	1	D70570	probable acyl-CoA
145	23	82.1	679	2	T20568	hypothetical prote	218	22	78.6	288	2	E70580	hypothetical prote
146	23	82.1	693	2	G82618	pius biogenesis p	219	22	78.6	293	2	AC2780	tolB protein (lipo
147	23	82.1	695	2	T02630	hypothetical prote	220	22	78.6	293	2	E97559	tolB protein precu
148	23	82.1	702	2	T52634	hypothetical prote	221	22	78.6	295	2	S23585	carboxyphosphoen
149	23	82.1	707	2	T32410	nematode responsiv	222	22	78.6	297	2	G69525	formylmethanofuran
150	23	82.1	723	2	I39066	hypothetical prote	223	22	78.6	298	2	C69819	ATP-dependent prot
151	23	82.1	737	2	T46040	N-methyl-D-asparta	224	22	78.6	300	2	AH3386	ABC transporter (A
152	23	82.1	739	2	S47772	biotin sulfoxide r	225	22	78.6	300	2	AE3328	hflC protein (EC 3
153	23	82.1	739	2	D91183	biotin sulfoxide r	226	22	78.6	305	2	AF0305	sugar binding prot
154	23	82.1	739	2	H86029	biotin sulfoxide r	227	22	78.6	306	2	AG2698	cation efflux syst
155	23	82.1	822	2	T33163	hypothetical prote	228	22	78.6	306	2	G97480	hypothetical prote
156	23	82.1	888	2	T31131	pyruvate, phosphat	229	22	78.6	308	2	E98259	hypothetical prote
157	23	82.1	909	2	H86350	hypothetical prote	230	22	78.6	308	2	AD3025	transcription regu
158	23	82.1	1008	2	T18832	probable RNA helic	231	22	78.6	309	2	A42921	transcription elon
159	23	82.1	1071	2	S38164	ATP-binding protei	232	22	78.6	321	2	US0772	37K glycoprotein -
160	23	82.1	1312	2	T30845	probable DNA repai	233	22	78.6	332	2	F82140	C4-dicarboxylate-b
161	23	82.1	1314	2	T47331	hypothetical prote	234	22	78.6	334	2	A10795	probable receptor/
162	23	82.1	1385	2	T21706	hypothetical prote	235	22	78.6	339	2	G83900	hypothetical prote
163	23	82.1	2004	2	AC0314	probable membrane	236	22	78.6	344	2	D86362	hypothetical prote
164	23	82.1	2535	2	AC0304	probable hemolysin	237	22	78.6	349	2	T50372	probable ATP-depen
165	23	82.1	6420	2	AC2981	polyketide synthas	238	22	78.6	355	2	T36273	hypothetical prote
166	23	82.1	60	2	AC2981	hypothetical prote	239	22	78.6	366	2	T16119	hypothetical prote
167	23	78.6	94	2	S69114	fibrinogen alpha c	240	22	78.6	374	2	C87134	conserved hypoteth
168	23	78.6	99	2	S74615	hypothetical prote	241	22	78.6	376	2	A75360	hypothetical prote
169	23	78.6	100	2	T45169	hypothetical prote	242	22	78.6	377	2	S77634	exopolysaccharide
170	23	78.6	103	2	G87367	flagellar hook-bas	243	22	78.6	379	2	D70846	probable metaA prot
171	23	78.6	116	2	D81945	probable pilus bio	244	22	78.6	381	2	C70711	hypothetical prote
172	23	78.6	116	2	G81159	probable type IV p	245	22	78.6	381	2	A23357	hypothetical prote
173	23	78.6	123	2	B45355	hypothetical prote	246	22	78.6	382	2	T45301	homoserine o-acety
174	23	78.6	123	2	S64248	hypothetical prote	247	22	78.6	386	2	T09019	phosphoprotein pho
175	23	78.6	126	2	S74493	thioredoxin M-2 -	248	22	78.6	386	2	A96532	hypothetical prote

249	22	78.6	388	2	C85323	protein phosphatas	322	78.6	875	2	F70755	hypothetical prote
250	22	78.6	389	2	F82809	conserved hypoteth	323	78.6	883	2	A71434	probable RNA helic
251	22	78.6	389	2	G95939	probable hippurate	324	78.6	884	2	C70729	hypothetical prote
252	22	78.6	391	2	AG0533	probable drug effl	325	78.6	901	1	RGEQWT	regulatory protein
253	22	78.6	398	2	JC2516	formaldehyde diamu	326	78.6	901	1	AG0996	Malt regulatory pr
254	22	78.6	398	2	S29819	serpin - barley	327	78.6	901	2	D91161	positive regulator
255	22	78.6	401	2	T15607	hypothetical prote	328	78.6	901	2	D86007	positive regulator
256	22	78.6	403	2	D82072	tRNA nucleotidyltr	329	78.6	903	2	AI0015	maltose regulon po
257	22	78.6	403	2	D96961	carboxyl-terminal	330	78.6	956	1	A46016	thrombospondin 3 -
258	22	78.6	409	2	B25156	cellulase (EC 3.2.	331	78.6	956	1	G70327	isoleucine-CRNA li
259	22	78.6	410	2	A45457	alpha 1-proteinase	332	78.6	956	2	A57121	thrombospondin 3 p
260	22	78.6	411	2	H83788	involved in spore	333	78.6	968	2	T46568	ATP-dependent RNA
261	22	78.6	416	2	B64132	tRNA adenyllytrans	334	78.6	1003	2	T24545	hypothetical prote
262	22	78.6	416	2	H95920	probable membrane-	335	78.6	1052	2	H83909	cell wall-associat
263	22	78.6	417	2	A55524	aspartic proteinas	336	78.6	1060	2	T30347	vitellogenin conve
264	22	78.6	418	2	S42072	aspartic proteinas	337	78.6	1078	2	T30879	dyein heavy chain
265	22	78.6	419	2	T10652	hypothetical prote	338	78.6	1087	2	D84767	probable pre-mRNA
266	22	78.6	420	2	T08691	hypothetical prote	339	78.6	1090	2	C86450	F5D14.27 protein -
267	22	78.6	424	2	T01383	GTPase-activating	340	78.6	1386	2	T49316	profilaggrin relat
268	22	78.6	429	2	AH1009	sun protein [impor	341	78.6	1599	2	T15854	chemotaxis protein
269	22	78.6	431	2	AG1884	glycolate oxidase	342	78.6	1638	2	E71410	probable centromer
270	22	78.6	436	2	T22253	hypothetical prote	343	78.6	1676	2	T17465	rifamycin polyketi
271	22	78.6	454	2	G81828	probable periplasm	344	78.6	2051	2	T13164	phloxin B - fruit f
272	22	78.6	456	2	E83609	probable aminotran	345	78.6	2116	2	A26655	myosin heavy chain
273	22	78.6	458	2	E69528	cobyrinic acid a,c	346	78.6	2166	2	G70163	hypothetical prote
274	22	78.6	458	2	G83690	hypothetical prote	347	78.6	2472	2	E83594	still frameshift p
275	22	78.6	459	2	S33424	repsA protein - St	348	78.6	2475	2	T00047	gellan lyase (EC 4
276	22	78.6	466	2	B98186	amidohydrolase [im	349	78.6	2695	2	S54974	type 1 inositol 1,
277	22	78.6	466	2	D69716	probable hydrolase	350	78.6	2713	2	A55713	inositol 1,4,5-tri
278	22	78.6	468	2	D69716	involved in spore	351	78.6	2734	2	B36579	inositol 1,4,5-tri
279	22	78.6	473	2	H84550	probable obtusifol	352	78.6	2749	1	ACMSIT	inositol 1,4,5-tri
280	22	78.6	475	2	B81047	secretion protein,	353	78.6	2749	2	T17467	inositol 1,4,5-tri
281	22	78.6	484	1	B69056	IMP dehydrogenase	354	78.6	3413	2	A36579	rifamycin polyketi
282	22	78.6	486	2	H82377	ABC transporter, A	355	78.6	5069	2	T17464	gp66.1 protein - M
283	22	78.6	487	2	T22825	hypothetical prote	356	78.6	46	2	B72808	conserved hypoteth
284	22	78.6	488	2	D86250	hypothetical prote	357	75.0	73	2	H69460	kcrA1 protein - pl
285	22	78.6	492	2	T30011	hypothetical prote	358	75.0	77	2	S00970	kleA protein - Ent
286	22	78.6	505	2	JC1458	trans-cinnamate 4-	359	75.0	80	2	T08483	hypothetical prote
287	22	78.6	508	1	O4CHC7	steroid 17alpha-mo	360	75.0	82	2	T49980	hypothetical prote
288	22	78.6	510	2	B68116	4-aminobutyrate tr	361	75.0	82	2	G97763	hypothetical prote
289	22	78.6	517	2	T10857	trans-cinnamate 4-	362	75.0	83	2	S76010	hypothetical prote
290	22	78.6	517	2	D75494	cell division prot	363	75.0	88	1	EDBE51	immediate-early-5
291	22	78.6	520	2	C87450	GMP synthase [impo	364	75.0	95	2	D83360	hypothetical prote
292	22	78.6	521	2	C87474	hypothetical prote	365	75.0	99	2	A69514	hypothetical prote
293	22	78.6	533	2	AG2293	hypothetical prote	366	75.0	102	1	R3HS15	ribosomal protein
294	22	78.6	548	2	G82286	phosphate ABC tran	367	75.0	104	2	T00845	probable retroelem
295	22	78.6	553	2	D95071	metallo-beta-lacta	368	75.0	106	2	T47202	H+-transporting tw
296	22	78.6	553	2	B97939	conserved hypoteth	369	75.0	109	2	E72850	AcOrf1-5 protein -
297	22	78.6	555	2	AB1254	hypothetical prote	370	75.0	109	2	T41895	orf5 homolog orf13
298	22	78.6	555	2	AH1616	hypothetical prote	371	75.0	110	2	H85234	hypothetical prote
299	22	78.6	557	2	C83130	probable sulfite o	372	75.0	110	2	H85234	hypothetical prote
300	22	78.6	573	2	T25397	hypothetical prote	373	75.0	112	2	F70403	H+-transporting tw
301	22	78.6	575	1	WQBEP1	phosphotransferase	374	75.0	113	2	A36493	hypothetical prote
302	22	78.6	584	2	E83767	ABC transporter (A	375	75.0	115	2	E97483	hypothetical prote
303	22	78.6	593	2	S26696	alkaline proteinas	376	75.0	120	2	B84364	H+-transporting Ar
304	22	78.6	624	1	BHTLE	hemocyanin chain e	377	75.0	129	2	AG0127	conserved hypoteth
305	22	78.6	637	2	G84356	hypothetical prote	378	75.0	130	2	A53641	arsenate reductase
306	22	78.6	655	2	E82712	peptidyl-prolyl ci	379	75.0	131	1	C41902	arsenate reductase
307	22	78.6	660	2	T20569	hypothetical prote	380	75.0	131	1	D41903	arsenate reductase
308	22	78.6	662	2	T20570	hypothetical prote	381	75.0	131	1	AF1944	homeotic protein L
309	22	78.6	687	2	T08528	probable DNA topoi	382	75.0	132	2	S43488	hypothetical prote
310	22	78.6	699	2	S67773	hypothetical prote	383	75.0	136	2	T22240	hypothetical prote
311	22	78.6	714	2	E81751	transcription elon	384	75.0	138	1	QQVLS	gene X protein - g
312	22	78.6	714	2	AF2458	hypothetical prote	385	75.0	138	2	H75120	translation initia
313	22	78.6	735	2	S37903	probable ATP-depen	386	75.0	138	2	G71010	arsenate reductase
314	22	78.6	755	2	S58718	probable nuclear p	387	75.0	139	2	F84024	arsenate reductase
315	22	78.6	768	2	AE3505	alkaline phosphata	388	75.0	140	2	A84284	hypothetical prote
316	22	78.6	797	2	S53590	hypothetical prote	389	75.0	141	1	QQVLC1	gene X protein - w
317	22	78.6	804	2	H75549	glycogen debranchi	390	75.0	141	2	AH2767	conserved hypoteth
318	22	78.6	826	2	T01494	trehalose-6-phosph	391	75.0	141	2	C97548	hypothetical prote
319	22	78.6	872	2	B86602	alanyl tRNA synthet	392	75.0	142	2	B70360	arsenate reductase
320	22	78.6	872	2	H72023	alanine-tRNA ligas	393	75.0	146	2	T08344	hypothetical prote
321	22	78.6	875	2	S27572	inner-membrane pro	394	75.0				

395 21 75.0 146 2 E90556 hypothetical prote
396 21 75.0 148 2 D64386 ribosomal protein
397 21 75.0 149 2 AF2977 conserved hypochet
398 21 75.0 150 2 A83217 conserved hypochet
399 21 75.0 160 2 F98305 hypothetical 16.6K
400 21 75.0 161 2 S66152 probable DNA-direc
401 21 75.0 163 2 C85582 unknown protein en
402 21 75.0 163 2 F90731 hypothetical prote
403 21 75.0 164 2 T30991 hypothetical prote
404 21 75.0 165 2 T47011 hypothetical prote
405 21 75.0 165 2 AF0237 probable thiorodox
406 21 75.0 167 2 E81028 conserved hypochet
407 21 75.0 167 2 A81973 hypothetical prote
408 21 75.0 167 2 AE0300 conserved hypochet
409 21 75.0 168 2 B84211 hypothetical prote
410 21 75.0 172 2 AG1103 B. subtilis yach p
411 21 75.0 178 2 S62846 H+-transporting tw
412 21 75.0 179 2 H64093 ribosomal protein
413 21 75.0 180 2 A69387 fumate (fum-1) h
414 21 75.0 180 2 F75607 arsenate reductase
415 21 75.0 182 2 T64029 hypothetical prote
416 21 75.0 183 1 TVHUR2 transforming prote
417 21 75.0 183 2 S03180 transforming prote
418 21 75.0 186 2 T46085 proto-oncogene - c
419 21 75.0 188 2 T19507 hypothetical prote
420 21 75.0 188 2 D89819 conserved hypochet
421 21 75.0 191 2 S09635 pabA protein - Ser
422 21 75.0 197 2 AF1742 protein involved i
423 21 75.0 197 2 AH1372 proteins involved
424 21 75.0 197 2 AC3447 outer membrane pro
425 21 75.0 200 2 T42678 hypothetical prote
426 21 75.0 201 2 AL3639 hypothetical prote
427 21 75.0 202 2 B83761 hypothetical prote
428 21 75.0 204 2 JC4261 somatotropin precu
429 21 75.0 206 2 AC1823 hypothetical prote
430 21 75.0 207 2 F87470 transcription regu
431 21 75.0 211 2 T42603 gene 60 protein -
432 21 75.0 212 1 WZBEF2 gene 60 protein -
433 21 75.0 212 2 S56199 hypothetical prote
434 21 75.0 214 2 E83431 type III export pr
435 21 75.0 214 2 A47715 cardiac-specific h
436 21 75.0 219 2 AC3026 transcription regu
437 21 75.0 222 2 S77550 hypothetical prote
438 21 75.0 224 2 D75150 hypothetical prote
439 21 75.0 224 2 D70685 probable urase ac
440 21 75.0 226 2 AB1741 ABC-transporter AT
441 21 75.0 226 2 AD1371 ABC-transporter AT
442 21 75.0 229 2 H75594 hypothetical prote
443 21 75.0 232 2 C69882 glycine betaine/L-
444 21 75.0 233 2 S23522 transcription init
445 21 75.0 234 2 F95127 hypothetical prote
446 21 75.0 234 2 C97998 conserved hypochet
447 21 75.0 237 2 AF1566 2-C-methyl-D-eryth
448 21 75.0 237 2 G84299 hypothetical prote
449 21 75.0 238 2 F98258 hypothetical trans
450 21 75.0 238 2 T29797 hypothetical prote
451 21 75.0 242 2 S69450 spermidine/putresc
452 21 75.0 242 2 D90796 hypothetical prote
453 21 75.0 243 2 T26653 hypothetical prote
454 21 75.0 244 2 G82266 pseudouridine synt
455 21 75.0 244 2 D86275 F7A19.26 protein -
456 21 75.0 245 2 AC2540 two-component resp
457 21 75.0 246 2 T36203 probable integral
458 21 75.0 252 2 S74562 ABC-type transport
459 21 75.0 254 2 T31709 hypothetical prote
460 21 75.0 256 2 A72352 hypothetical prote
461 21 75.0 257 2 B89124 protein K07C11.1
462 21 75.0 257 2 B75099 hypothetical prote
463 21 75.0 258 2 T34601 hypothetical prote
464 21 75.0 259 2 AB0565 probable membrane
465 21 75.0 262 2 S63648 H+-transporting tw
466 21 75.0 262 2 T18886 hypothetical prote
467 21 75.0 264 2 T26558 hypothetical prote

468 21 75.0 268 2 D85548 probable metal res
469 21 75.0 268 2 B90698 probable metal res
470 21 75.0 268 2 B64780 probable transport
471 21 75.0 268 2 T23665 hypothetical prote
472 21 75.0 268 2 AG1225 cobalt transport A
473 21 75.0 268 2 A11578 teichoic acid ABC
474 21 75.0 269 2 D86739 flagellin FljM [im
475 21 75.0 273 2 E87347 flagellin FljN [im
476 21 75.0 273 2 F87347 flagellin FljK [im
477 21 75.0 273 2 F87430 cysteine synthase
478 21 75.0 274 2 G84218 MBT8 protein - yea
479 21 75.0 274 2 S20155 conserved hypochet
480 21 75.0 275 2 C70924 methylesterase,
481 21 75.0 276 2 B87682 probable membrane
482 21 75.0 277 1 QOECAD probable beta-lact
483 21 75.0 277 2 G75518 conserved hypochet
484 21 75.0 277 2 AC0704 hypothetical prote
485 21 75.0 277 2 F85778 hypothetical prote
486 21 75.0 277 2 B90930 hypothetical prote
487 21 75.0 279 2 C75491 probable serine es
488 21 75.0 281 2 A84259 hypothetical prote
489 21 75.0 285 2 G85605 unknown in ISB88
490 21 75.0 287 2 S25259 streptomycin resis
491 21 75.0 288 2 S57184 aryl-alcohol dehyd
492 21 75.0 290 2 A87312 coproporphyrinogen
493 21 75.0 290 2 A84011 transcription acti
494 21 75.0 290 2 A69779 conserved hypochet
495 21 75.0 292 2 H64149 hypothetical prote
496 21 75.0 294 2 AB2367 dTDP-6-deoxy-L-man
497 21 75.0 294 2 C69956 phosphate ABC tran
498 21 75.0 297 2 A84767 hypothetical prote
499 21 75.0 297 2 S66102 protein secretion
500 21 75.0 298 2 D95199 primosomal protein
501 21 75.0 298 2 B98066 primosome componen
502 21 75.0 304 2 A96642 FMRI protein - Pod
503 21 75.0 305 2 S19797 probable LysR - typ
504 21 75.0 307 2 D95363 dioxygenase relate
505 21 75.0 310 2 H97338 oxidoreductase, al
506 21 75.0 314 2 D87376 bifunctional cycla
507 21 75.0 315 2 S58171 hypothetical prote
508 21 75.0 315 2 T10613 hypothetical prote
509 21 75.0 316 2 B90282 hypothetical prote
510 21 75.0 317 2 AI0640 flagellar hook-ass
511 21 75.0 317 2 S10362 hook-associated pr
512 21 75.0 318 2 D75210 hypothetical prote
513 21 75.0 324 2 T11571 acetyl-CoA carboxy
514 21 75.0 325 2 A75330 thiorodoxin reduct
515 21 75.0 325 2 T48873 electron transfer
516 21 75.0 327 2 F85708 unknown protein en
517 21 75.0 329 2 S67807 hypothetical prote
518 21 75.0 330 2 A82538 pseudouridylylate sy
519 21 75.0 331 2 AF3467 peptidyl-prolyl ci
520 21 75.0 333 2 T48363 histidyl-trNA synt
521 21 75.0 333 2 D83585 hypothetical prote
522 21 75.0 334 2 F82149 Holliday junction
523 21 75.0 334 2 T19860 hypothetical prote
524 21 75.0 335 2 T07517 DNA-directed RNA p
525 21 75.0 338 2 S67158 26S proteasome reg
526 21 75.0 340 2 AC0459 porphobilinogen sy
527 21 75.0 341 2 B87390 transcription regu
528 21 75.0 342 2 T08474 kfrA protein - Ent
529 21 75.0 343 2 S74937 hypothetical prote
530 21 75.0 346 1 RQPSNA RacA protein PA361
531 21 75.0 346 2 C81446 homolog of E. coli
532 21 75.0 346 2 H95406 conserved hypochet
533 21 75.0 347 2 B64691 rod shape-determin
534 21 75.0 347 2 B71826 rod shape-determin
535 21 75.0 347 2 S22617 hypothetical prote
536 21 75.0 348 2 C87313 transcription regu
537 21 75.0 348 2 H90850 hypothetical prote
538 21 75.0 349 2 E84400 dithiorotate deh
539 21 75.0 349 2 JN0321 recombination prot
540 21 75.0 349 2 S22448 FMRI protein - Pod

541	21	75.0	349	2	H85626	hypothetical prote	614	21	75.0	434	2	F75425	tRNA nucleotidyltr
542	21	75.0	349	2	B30819	hypothetical prote	615	21	75.0	435	2	T45974	phenylalanine-tRNA
543	21	75.0	349	2	E80912	hypothetical prote	616	21	75.0	435	2	A84828	hypothetical prote
544	21	75.0	350	2	B88691	protein F41H10.10	617	21	75.0	436	2	A11316	hypothetical prote
545	21	75.0	355	2	F81249	peroxidase (EC 1.1	618	21	75.0	436	2	A11688	hypothetical prote
546	21	75.0	359	2	F83850	hypothetical prote	619	21	75.0	438	2	JC1179	transforming prote
547	21	75.0	361	2	G95410	probable ABC trans	620	21	75.0	439	1	TVCTMC	transforming prote
548	21	75.0	362	2	S19419	hypothetical prote	621	21	75.0	439	1	TVHUM	transforming prote
549	21	75.0	363	2	S14355	peroxidase (EC 1.1	622	21	75.0	439	1	TVMS	transforming prote
550	21	75.0	364	2	S34355	probable potassium	623	21	75.0	439	1	TVTMC	transforming prote
551	21	75.0	364	2	T43361	hypothetical 43.5K	624	21	75.0	439	1	TVTMC	transforming prote
552	21	75.0	365	2	JU0319	hypothetical prote	625	21	75.0	439	2	JU0449	transforming prote
553	21	75.0	368	2	A11934	hypothetical prote	626	21	75.0	440	4	TVHUT	probable pyridine
554	21	75.0	370	2	F70484	twitching mobility	627	21	75.0	441	2	AD0572	T24D18.11 protein
555	21	75.0	370	2	AC1275	aminopeptidase hom	628	21	75.0	442	2	G86294	hypothetical prote
556	21	75.0	370	2	AC1638	hypothetical prote	629	21	75.0	442	1	QQBYPT	alpha-amylase - Ae
557	21	75.0	370	2	A11143	hypothetical prote	630	21	75.0	443	1	I39538	hypothetical prote
558	21	75.0	373	1	KMECTD	chiasmata mutase	631	21	75.0	443	2	H96841	hypothetical prote
559	21	75.0	373	1	S29934	chiasmata mutase	632	21	75.0	443	2	C83421	sigma-54 dependant
560	21	75.0	373	2	G91061	chiasmata mutase-	633	21	75.0	444	2	B82296	glucose inhibited
561	21	75.0	373	2	C85906	chiasmata mutase-	634	21	75.0	444	2	D97977	hypothetical prote
562	21	75.0	373	2	A80399	prephenate dehydro	635	21	75.0	445	2	E72674	hypothetical prote
563	21	75.0	373	2	A10832	prephenate dehydro	636	21	75.0	445	2	A83763	L-arabinose transp
564	21	75.0	375	2	S47455	mitosis protein JN	637	21	75.0	447	2	G95068	cysteine-tRNA syn
565	21	75.0	375	2	E82292	chiasmata mutase/	638	21	75.0	447	2	G97936	cysteine-tRNA liga
566	21	75.0	376	2	S11335	probable aryl-alco	639	21	75.0	448	2	C96542	hypothetical prote
567	21	75.0	376	2	T10594	hypothetical prote	640	21	75.0	449	2	H70526	probable cytochrom
568	21	75.0	377	2	D86881	pyridine nucleotid	641	21	75.0	450	2	H64756	probable mercury(I
569	21	75.0	378	2	E84806	probable elongatid	642	21	75.0	450	2	B85522	probable oxidoredu
570	21	75.0	379	1	JC4289	cell division prot	643	21	75.0	450	2	F90671	probable membrane
571	21	75.0	380	1	E69541	conserved hypotet	644	21	75.0	451	2	AG0008	gag-myc polyprotei
572	21	75.0	381	2	T46827	phenoxycarboxate di	645	21	75.0	451	1	TVFV2C	aryl hydrocarbon r
573	21	75.0	384	2	H89873	hypothetical prote	646	21	75.0	451	2	T42397	Similar to Flavono
574	21	75.0	385	2	H87355	peptide chain rele	647	21	75.0	452	2	F96672	two component sens
575	21	75.0	385	2	H95413	probable integrase	648	21	75.0	452	2	AC3136	beta-alanine-pyruv
576	21	75.0	387	2	AC0887	probable alcohol d	649	21	75.0	452	2	AH3525	hypothetical prote
577	21	75.0	387	2	T09086	sedoheptulose-bisp	650	21	75.0	453	2	T19440	conserved hypotet
578	21	75.0	387	2	A33497	prophage DLP12 int	651	21	75.0	456	1	E71284	Gid protein (impor
579	21	75.0	389	2	A11506	conserved hypotet	652	21	75.0	456	1	A95109	t1d4D-related prote
580	21	75.0	389	2	H69048	conserved hypotet	653	21	75.0	458	1	E71033	argininosuccinate
581	21	75.0	390	2	A11147	hypothetical prote	654	21	75.0	458	2	B81175	probable argininos
582	21	75.0	396	2	B82320	conserved hypotet	655	21	75.0	460	2	C81930	cysteine-tRNA syn
583	21	75.0	398	2	S65782	membrane-bound lyt	656	21	75.0	460	2	G83421	hypothetical prote
584	21	75.0	399	2	T06488	serpin - whea	657	21	75.0	461	2	AH2291	cytochrome P450 ho
585	21	75.0	399	2	A25470	cytokeratin 19 - b	658	21	75.0	461	2	T07859	arginosuccinate
586	21	75.0	401	2	D87258	fatty oxidation co	659	21	75.0	464	2	C82756	peptidyl-prolyl ci
587	21	75.0	401	2	H87411	acyl-CoA dehydroge	660	21	75.0	464	2	S03325	transforming prote
588	21	75.0	405	2	T47595	RING finger protei	661	21	75.0	465	2	H84198	proline-tRNA synch
589	21	75.0	405	2	T41712	hypothetical prote	662	21	75.0	465	2	C83421	cysteine-tRNA liga
590	21	75.0	405	2	AB0793	probable MR-MLE-fa	663	21	75.0	466	1	C53402	poly(A) polymerase
591	21	75.0	405	2	G91020	probable racemase	664	21	75.0	466	2	G83663	cysteine-tRNA syn
592	21	75.0	405	2	T49534	hypothetical prote	665	21	75.0	467	2	D83055	hypothetical prote
593	21	75.0	405	2	F85864	probable racemase	666	21	75.0	468	2	G89912	hypothetical prote
594	21	75.0	405	2	H97130	ftsk-like DNA seg	667	21	75.0	468	2	C71336	glutamyl-tRNA (gln)
595	21	75.0	405	2	E64995	hypothetical prote	668	21	75.0	471	2	AB3422	probable monoxyme
596	21	75.0	407	2	G95054	hypothetical prote	669	21	75.0	472	2	G70932	probable polynucle
597	21	75.0	407	2	D97924	ROK family protein	670	21	75.0	472	2	H71345	probable cytochrom
598	21	75.0	410	2	T36535	xylose repressor p	671	21	75.0	481	2	B96691	gag-myc polyprotei
599	21	75.0	410	2	T36535	probable hydrolase	672	21	75.0	481	2	TVMVFT	flagellar hook-ass
600	21	75.0	411	2	T51818	ATP phosphoribosyl	673	21	75.0	484	1	AF3529	histidine-tRNA lig
601	21	75.0	413	2	T47272	transposase tnpA l	674	21	75.0	484	2	T49251	phosphoserine phos
602	21	75.0	416	1	FOCH	transforming prote	675	21	75.0	486	2	C69084	head protein gp5 -
603	21	75.0	416	1	TVFVAC	transforming prote	676	21	75.0	495	2	JN0539	probable head-tail
604	21	75.0	416	2	C96725	hypothetical prote	677	21	75.0	501	1	C85743	probable head-tail
605	21	75.0	423	2	G95880	transforming prote	678	21	75.0	501	2	D90767	probable head-tail
606	21	75.0	423	2	A98152	probable trehalose	679	21	75.0	501	2	D90970	hypothetical prote
607	21	75.0	428	2	AC2288	sensor histidine k	680	21	75.0	501	2	D85717	spliceosome-associ
608	21	75.0	432	2	A10611	cell division prot	681	21	75.0	501	2	A55749	probable carboxyle
609	21	75.0	432	4	A46932	hypothetical cell-	682	21	75.0	502	2	T35910	lipopolysaccharide
610	21	75.0	430	2	AE0168	seryl-tRNA synthet	683	21	75.0	502	2	A83938	pollen-specific pr
611	21	75.0	430	2	D83977	hypothetical prote	684	21	75.0	504	2	T07129	unknown protein, 7
612	21	75.0	431	2	F89817	conserved hypotet	685	21	75.0	508	2	F86458	
613	21	75.0	431	2	H71172	hypothetical prote	686	21	75.0	508	2		

687	21	75.0	508	2	T22636	hypothetical prote	760	21	75.0	668	2	AG1989	hypothetical prote
688	21	75.0	509	2	T00793	hypothetical prote	761	21	75.0	670	2	S53414	probable membrane
689	21	75.0	525	2	S69991	alcohol O-acetyltr	762	21	75.0	673	1	BVECTB	excinuclease ABC c
690	21	75.0	525	2	S67289	alcohol O-acetyltr	763	21	75.0	673	2	AD0597	excision nuclease
691	21	75.0	526	2	S54035	Gpase MSS1, mitoc	764	21	75.0	673	2	B8586	DNA repair, excisi
692	21	75.0	528	2	G97736	hypothetical prote	765	21	75.0	673	2	A99736	excision nuclease
693	21	75.0	529	2	T10612	hypothetical prote	766	21	75.0	678	2	H8187	protein Cl8H9.8 [i
694	21	75.0	535	2	C95057	Ctp synthase [lipo	767	21	75.0	679	2	G64112	excinuclease ABC c
695	21	75.0	535	2	F97926	Ctp synthase (BC 6	768	21	75.0	680	2	D89625	protein Cl4H10.2 [
696	21	75.0	538	2	T08642	hypothetical prote	769	21	75.0	684	2	H96846	hypothetical prote
697	21	75.0	541	2	C96492	probable pectinest	770	21	75.0	684	2	T02149	hypothetical prote
698	21	75.0	541	2	T49108	pectinesterase lik	771	21	75.0	692	2	A70933	probable nrndz prot
699	21	75.0	542	2	S76358	GMP synthase (glut	772	21	75.0	693	2	C70167	translation elonga
700	21	75.0	547	2	G71307	probable oligopept	773	21	75.0	699	2	T16109	hypothetical prote
701	21	75.0	548	2	T05670	pollen-specific pr	774	21	75.0	699	2	T18426	hypothetical prote
702	21	75.0	548	2	A71544	probable diphospha	775	21	75.0	701	2	JC8062	L-glutamate oxidas
703	21	75.0	549	2	T02790	hypothetical prote	776	21	75.0	701	2	S35313	hypothetical prote
704	21	75.0	551	1	Q0V225	rifampicin resista	777	21	75.0	701	2	T28151	probable ABC-type
705	21	75.0	551	2	H86351	protein T26P17.6 [778	21	75.0	709	2	T00096	DNA replication-re
706	21	75.0	551	2	A36848	N3L protein - vari	779	21	75.0	712	2	T33028	hypothetical prote
707	21	75.0	551	2	T28541	hypothetical prote	780	21	75.0	719	2	A10769	protein-tyrosine k
708	21	75.0	551	2	B72163	O3L protein - vari	781	21	75.0	724	2	AB1199	ATP-dependent prot
709	21	75.0	552	2	T33760	sphingosine-1-phos	782	21	75.0	724	2	AC1557	ATP-dependent prot
710	21	75.0	554	2	S22495	pollen-specific pr	783	21	75.0	728	2	AB0101	probable exported
711	21	75.0	560	2	S54093	hypothetical prote	784	21	75.0	730	2	G75292	excinuclease ABC c
712	21	75.0	561	2	T05545	pollen-specific pr	785	21	75.0	734	2	AF3108	GGDPF family prote
713	21	75.0	565	2	T22649	hypothetical prote	786	21	75.0	737	2	C70770	hypothetical prote
714	21	75.0	569	1	S35291	urease (EC 3.5.1.5	787	21	75.0	739	2	F71161	probable chemotaxi
715	21	75.0	570	2	S52765	secD protein - Str	788	21	75.0	742	2	E89801	hypothetical prote
716	21	75.0	575	1	WQECPI	phosphotransferase	789	21	75.0	746	2	T19287	hypothetical prote
717	21	75.0	575	1	SVBYNM	methionine-tRNA li	790	21	75.0	747	2	F98178	nitrogen fixation
718	21	75.0	575	2	C85894	PRP-protein phosph	791	21	75.0	754	2	AE0614	probable competenc
719	21	75.0	575	2	H91039	PRP-protein phosph	792	21	75.0	758	2	F72363	hypothetical prote
720	21	75.0	577	2	A57469	hypothetical prote	793	21	75.0	761	2	T03719	probable thyridi r
721	21	75.0	580	2	T21493	hypothetical prote	794	21	75.0	775	1	WM8E19	ribonucleoside-dip
722	21	75.0	583	2	S01496	lamin B - African	795	21	75.0	786	2	E86268	Fi3B4.4 protein -
723	21	75.0	583	2	F70592	probable lpq8 prot	796	21	75.0	786	2	JC1298	aculeacin-A acylas
724	21	75.0	584	2	S05518	lamin B-1 - chicke	797	21	75.0	787	2	H86707	formate C-acetyltr
725	21	75.0	584	2	T23174	hypothetical prote	798	21	75.0	788	2	S48191	probable ubiquinol
726	21	75.0	586	1	VEHUJ4	lamin B1 - human	799	21	75.0	807	2	E90523	leucyl-trna synth
727	21	75.0	587	2	T07634	pollen-specific pr	800	21	75.0	811	2	T36581	probable transmem
728	21	75.0	587	2	AD3519	2,3-dihydroxybenzo	801	21	75.0	813	2	T46242	kinesin-like prote
729	21	75.0	587	2	S07720	lamin B - mouse	802	21	75.0	817	2	T04697	hypothetical prote
730	21	75.0	597	2	D70100	phosphoglucosylase	803	21	75.0	822	2	S77112	DNA mismatch repai
731	21	75.0	598	1	A37192	excinuclease ABC,	804	21	75.0	830	2	T16236	hypothetical prote
732	21	75.0	600	2	T36195	probable acyl-CoA	805	21	75.0	844	2	S61104	BR01 protein - yea
733	21	75.0	601	2	T49752	hypothetical prote	806	21	75.0	846	2	JC7721	aryl hydrocarbon r
734	21	75.0	603	2	T38920	methylenetetrahydr	807	21	75.0	862	2	E84567	probable trehalose
735	21	75.0	606	2	F69493	conserved hypothet	808	21	75.0	870	2	S65158	PAL1 protein - yea
736	21	75.0	609	1	A42537	gene 16 protein -	809	21	75.0	871	2	C72238	preprotein translo
737	21	75.0	616	2	G86890	proline-RNA ligas	810	21	75.0	872	2	JC7380	DNA-directed DNA p
738	21	75.0	619	2	S25334	H-exporting Atpas	811	21	75.0	873	2	E90581	hypothetical prote
739	21	75.0	622	2	JC5425	transcription init	812	21	75.0	875	1	FOFVHB	sag-myc polyprotei
740	21	75.0	629	2	E82624	thiamin biosynthes	813	21	75.0	875	1	FOFVHB	sag-myc polyprotei
741	21	75.0	629	2	T06675	hypothetical prote	814	21	75.0	877	2	T35861	probable large sec
742	21	75.0	630	2	T38023	probable transrip	815	21	75.0	884	2	E71289	probable penicilli
743	21	75.0	631	2	A53623	yolk protein facto	816	21	75.0	887	2	B96598	hypothetical prote
744	21	75.0	636	2	F69037	cleavage and polya	817	21	75.0	890	2	E81576	translation initia
745	21	75.0	641	2	T49477	phenol hydroxylase	818	21	75.0	890	2	E86530	translation initia
746	21	75.0	643	2	T19199	hypothetical prote	819	21	75.0	890	2	F72093	translation initia
747	21	75.0	644	2	AD3471	adenylyl-sulfate k	820	21	75.0	896	2	G81709	isoleucyl-tRNA syn
748	21	75.0	645	2	S50374	GTP-binding protei	821	21	75.0	900	2	G82894	hypothetical prote
749	21	75.0	651	2	E86242	hypothetical prote	822	21	75.0	903	2	T47316	two-component sens
750	21	75.0	652	2	G82401	methyl-accepting c	823	21	75.0	911	2	AD0959	hypothetical prote
751	21	75.0	653	2	H97777	DNA helicase II (E	824	21	75.0	925	2	T29311	glutamate-ammonia
752	21	75.0	653	2	S67035	probable membrane	825	21	75.0	948	2	C82077	excinuclease ABC (
753	21	75.0	654	2	I56134	tumor necrosis fac	826	21	75.0	957	2	B84099	glutamate-ammonia
754	21	75.0	658	2	F71703	DNA helicase II (u	827	21	75.0	960	2	B84056	DNA translocase (s
755	21	75.0	658	2	T05655	hypothetical prote	828	21	75.0	966	2	D96714	DNA-directed RNA p
756	21	75.0	659	1	A64228	DNA ligase (NAD) (829	21	75.0	982	2	T04542	hypothetical prote
757	21	75.0	661	2	C83843	hypothetical prote	830	21	75.0	990	2	B49351	bacteriophage N4 a
758	21	75.0	661	2	T15802	hypothetical prote	831	21	75.0	990	2	H80703	bacteriophage N4 a
759	21	75.0	663	2	B70460	excinuclease ABC c	832	21	75.0	990	2	C85554	bacteriophage N4 a

833	21	75.0	991	2	A99334	hypothetical prote	906	20	71.4	85	2	D87324	hypothetical prote
834	21	75.0	1011	2	T51399	DNA-directed RNA p	907	20	71.4	87	2	E69932	hypothetical prote
835	21	75.0	1013	2	T30818	hesR protein - Kle	908	20	71.4	90	2	T42122	probable transposa
836	21	75.0	1019	2	A38738	coagulation factor	909	20	71.4	93	2	A95121	hypothetical prote
837	21	75.0	1021	2	T05901	hypothetical prote	910	20	71.4	93	2	F97990	hypothetical prote
838	21	75.0	1023	2	A30203	AcR3/AcrD/AcrF fam	911	20	71.4	95	2	T45336	hypothetical prote
839	21	75.0	1023	2	B98261	probable rmd efflu	912	20	71.4	95	2	AB3126	hypothetical prote
840	21	75.0	1031	2	H81288	probable sugar tra	913	20	71.4	96	1	BUKV	basic blue protein
841	21	75.0	1045	2	B30239	hydroxymethylgluta	914	20	71.4	96	2	AH0115	conserved hypothet
842	21	75.0	1085	2	JC2227	probable helicase	915	20	71.4	96	2	D83228	hypothetical prote
843	21	75.0	1088	1	PIXRPR	inner layer protei	916	20	71.4	97	2	E86682	prophage p12 prote
844	21	75.0	1088	2	S39261	VPI protein - porc	917	20	71.4	97	2	A86754	methyitransferase
845	21	75.0	1100	2	G83376	probable trehalose	918	20	71.4	98	2	AD1146	probable ribosomal
846	21	75.0	1143	2	S46122	SNF2 protein homol	919	20	71.4	101	2	C83539	morphogene protein
847	21	75.0	1145	2	T05573	hypothetical prote	920	20	71.4	101	2	S75079	hypothetical prote
848	21	75.0	1148	2	F69685	pyruvate carboxyla	921	20	71.4	102	2	H86160	30S ribosomal prot
849	21	75.0	1170	2	A72287	hypothetical prote	922	20	71.4	102	2	B84355	periplasmic divale
850	21	75.0	1172	2	AD0438	probable exported	923	20	71.4	102	2	A69497	hypothetical prote
851	21	75.0	1204	2	B81947	probable exodeoxyr	925	20	71.4	103	2	D97247	hypothetical prote
852	21	75.0	1207	2	C70013	conserved hypothet	926	20	71.4	104	2	S12188	Bola protein [limpo
853	21	75.0	1226	2	T49915	pre-mRNA splicing	927	20	71.4	104	2	D97247	hypothetical prote
854	21	75.0	1235	2	S24109	phosphorylase kina	928	20	71.4	105	2	A10557	hypothetical prote
855	21	75.0	1236	2	E70977	hypothetical prote	929	20	71.4	106	2	AH0383	Bola protein [limpo
856	21	75.0	1277	2	S70306	hypothetical prote	930	20	71.4	106	2	G82795	hypothetical prote
857	21	75.0	1309	1	BVB9D9	RAD9 protein - yea	931	20	71.4	110	2	S29724	lignin peroxidase
858	21	75.0	1321	2	T42228	P-glycoprotein sis	932	20	71.4	113	2	D70580	hypothetical prote
859	21	75.0	1321	2	T42842	bile salt transpor	933	20	71.4	114	2	JC5238	galactosylceramide
860	21	75.0	1383	2	T06091	hypothetical prote	934	20	71.4	115	2	AB0657	integrase (truncat
861	21	75.0	1394	2	S66876	ATP-dependent tran	935	20	71.4	115	2	B85883	probable transcrip
862	21	75.0	1396	2	G71529	DNA-directed RNA p	936	20	71.4	116	2	H45893	T-cell receptor al
863	21	75.0	1396	2	T10627	hypothetical prote	937	20	71.4	116	2	S17567	AQN-3 protein - pi
864	21	75.0	1449	2	B84426	hypothetical prote	938	20	71.4	116	2	S39434	spermadhesin AQN-3
865	21	75.0	1474	2	B85188	retrotransposon li	939	20	71.4	116	2	T51015	hypothetical prote
866	21	75.0	1479	2	A89793	hypothetical prote	940	20	71.4	117	2	H91038	hypothetical prote
867	21	75.0	1498	2	AF1082	B. subtilis YUKA p	941	20	71.4	117	2	I68524	ribosomal protein
868	21	75.0	1498	2	AG1439	B. subtilis YUKA p	942	20	71.4	118	2	G84302	hypothetical prote
869	21	75.0	1509	2	T19486	hypothetical prote	943	20	71.4	119	2	S53460	hypothetical prote
870	21	75.0	1525	2	T14961	hypothetical prote	944	20	71.4	120	2	E81229	hypothetical prote
871	21	75.0	1551	2	F86342	F9H16.4 protein -	945	20	71.4	120	2	D69833	hypothetical prote
872	21	75.0	1580	2	T26204	hypothetical prote	946	20	71.4	121	2	A41940	IG heavy chain V r
873	21	75.0	1581	2	T71636	hypothetical prote	947	20	71.4	123	2	G64219	ribosomal protein
874	21	75.0	1622	2	T45240	hypothetical prote	948	20	71.4	124	2	S62816	phosphate acceptor
875	21	75.0	1648	2	S61654	probable membrane	949	20	71.4	124	2	S20545	hypothetical prote
876	21	75.0	1758	2	S57015	hypothetical prote	950	20	71.4	124	2	S11769	hypothetical prote
877	21	75.0	1780	2	T17272	hypothetical prote	951	20	71.4	126	2	S35338	MLC1536.21c protei
878	21	75.0	1864	2	T18485	probable 1-phospha	952	20	71.4	126	2	T11008	transposase all269
879	21	75.0	1900	2	S45530	hypothetical prote	953	20	71.4	127	2	AF2142	transposase all440
880	21	75.0	2058	2	E71436	hypothetical prote	954	20	71.4	127	2	AI2407	transposase all481
881	21	75.0	2265	2	T26183	hypothetical prote	955	20	71.4	127	2	AF2360	transposase all443
882	21	75.0	2697	2	T25444	hypothetical prote	956	20	71.4	127	2	AE2450	transposase all001
883	21	75.0	2748	2	S57976	nuclear migration	957	20	71.4	127	2	AI1808	transposase all361
884	21	75.0	3206	1	GNVSPV	genome polyprotein	958	20	71.4	127	2	AC2257	hypothetical prote
885	21	75.0	3227	2	T37964	probable ubiquitin	959	20	71.4	128	2	E70547	lactoylglutathione
886	21	75.0	3228	2	T21381	hypothetical prote	960	20	71.4	128	2	B97170	hypothetical prote
887	21	75.0	3434	1	GNWTMV	genome polyprotein	961	20	71.4	129	2	S77545	antigen B [impor
888	21	75.0	3491	2	T43231	probable 6-deoxyer	962	20	71.4	129	2	AF1089	arsenate reductase
889	21	75.0	3591	1	S21010	Filamentous hemag	963	20	71.4	130	2	B69047	hypothetical prote
890	21	75.0	3635	2	T10053	laminin alpha 5 ch	964	20	71.4	131	2	T01293	hypothetical prote
891	21	75.0	4385	2	T29042	hypothetical prote	965	20	71.4	132	2	B31211	hypothetical prote
892	21	75.0	4967	2	S72269	ryanodine receptor	966	20	71.4	132	2	T02116	hypothetical prote
893	21	75.0	4969	2	A37113	hypothetical prote	967	20	71.4	132	2	B83299	hypothetical prote
894	21	75.0	7576	2	T17428	PK506 polyketide s	968	20	71.4	133	2	AD1875	probable small pro
895	20	71.4	15	2	S57577	T cell receptor V-	969	20	71.4	133	2	AD1875	conserved hypothet
896	20	71.4	27	2	A61412	methane monooxygen	970	20	71.4	134	2	F71354	gene lmdr1 protei
897	20	71.4	60	2	S87781	hypothetical prote	971	20	71.4	135	2	B87396	sperm motility inh
898	20	71.4	62	2	F57301	hypothetical prote	972	20	71.4	135	2	B44438	hypothetical prote
899	20	71.4	62	2	AI2130	hypothetical prote	973	20	71.4	136	2	S72508	hypothetical prote
900	20	71.4	68	2	T10328	hypothetical prote	974	20	71.4	136	2	G70540	probable heavy met
901	20	71.4	75	2	E70870	hypothetical prote	975	20	71.4	137	2	T36635	hypothetical prote
902	20	71.4	76	2	F83566	hypothetical prote	976	20	71.4	137	2	T17718	hypothetical prote
903	20	71.4	78	2	T35577	hypothetical prote	977	20	71.4	137	2	AI3194	hypothetical prote
904	20	71.4	83	2	S28120	gas-vesicle operon	978	20	71.4	138	2	AB0538	probable secreted
905	20	71.4	84	2	B69014	ferredoxin 2[4Fe-4	979	20	71.4	138	2	AB0538	probable secreted

979 20 71.4 138 2 C87389
980 20 71.4 138 2 A87192
981 20 71.4 140 2 B87622
982 20 71.4 141 2 S14259
983 20 71.4 142 1 E64794
984 20 71.4 142 2 E70356
985 20 71.4 142 2 F90709
986 20 71.4 142 2 B85560
987 20 71.4 142 2 H90468
988 20 71.4 146 2 AD0524
989 20 71.4 147 2 C64223
990 20 71.4 147 2 G70209
991 20 71.4 148 2 A26192
992 20 71.4 148 2 C71829
993 20 71.4 148 2 B64687
994 20 71.4 150 2 B71329
995 20 71.4 151 2 H88480
996 20 71.4 151 2 A69195
997 20 71.4 151 2 T19087
998 20 71.4 152 2 D72341
999 20 71.4 152 2 AF2680
1000 20 71.4 152 2 D97462

ALIGNMENTS

RESULT 1
T02022
hypothetical protein T9E19.1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C:Accession: T02022
R:Stromatt, C.; Johnson, D.; Levy, T.
submitted to the EMBL Data Library, November 1998
A:Description: The sequence of A. thaliana T9E19.
A:Reference number: Z14496
A:Accession: T02022
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-543 <STR>
A:Cross-references: UNIPROT:Q9ZSH5; UNIPARC:UPI000009C860; EMBL:AF104920; NID:g3859610;
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 198/3; 253/3; 295/3; 330/3; 364/3; 384/3
A:Note: T9E19.1

Query Match 96.4%; Score 27; DB 2; Length 543;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 69 VRDQIRAAAL 77

RESULT 2
T47381
hypothetical protein T5C2.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47381
R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24463
A:Accession: T47381
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1113 <OBE>
A:Cross-references: UNIPROT:Q09180; UNIPARC:UPI00000A5BDF; EMBL:AL1138664
A:Experimental source: cultivar Columbia; BAC clone T5C2
C:Genetics:

A:Map position: 3
A:Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 732/3; 767/3; 801/3; 816/3; 845/3; 9
A:Note: T5C2.90

Query Match 96.4%; Score 27; DB 2; Length 1113;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 488 VRDQIRAAAL 496

RESULT 3

E86402
hypothetical protein F28L5.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86402
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1198 <STO>
A:Cross-references: UNIPROT:Q9C6N9; UNIPARC:UPI00000A5394; GB:AE005172; NID:g10998930; P
C:Genetics:
A:Map position: 1

Query Match 96.4%; Score 27; DB 2; Length 1198;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 486 VRDQIRATL 494

RESULT 4

F86386
hypothetical protein F14G11.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86386
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86386
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1201 <STO>
A:Cross-references: UNIPROT:Q9C607; UNIPARC:UPI00000A1C16; GB:AE005172; NID:g11560181; P
C:Genetics:
A:Map position: 1

Query Match 96.4%; Score 27; DB 2; Length 1201;


```
Best Local Similarity 66.7%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

Qy 1 VRDQIXXXL 9
Db 455 VRDQIRAAAL 463

RESULT 5
H96559
hypothetical protein F5F19.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: H96559
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1305 <STO>
A:Cross-references: UNIPROT:Q9ZU21; UNIPARC:UPI00000AA103; GB:AE005173; NID:g4220449; PI
C:Genetics:
A:Gene: F5F19.8
A:Map position: 1

Query Match 96.4%; Score 27; DB 2; Length 1305;
Best Local Similarity 66.7%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

Qy 1 VRDQIXXXL 9
Db 491 VRDQIRAAAL 499

RESULT 6
A84470
hypothetical protein At2g05560 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84470
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84470
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1472 <STO>
A:Cross-references: UNIPROT:Q9SL11; UNIPARC:UPI00000A7470; GB:AE002093; NID:g4581168; PI
C:Genetics:
A:Gene: At2g05560
A:Map position: 2

Query Match 96.4%; Score 27; DB 2; Length 1472;
Best Local Similarity 66.7%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

Qy 1 VRDQIXXXL 9
Db 403 VRDQIRAAAL 411
```

```
RESULT 7
T02599
hypothetical protein At2g14770 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F26C24.9
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02599; C84521
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A:Reference number: Z14680
A:Accession: T02599
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1756 <ROU>
A:Cross-references: UNIPROT:O80975; UNIPARC:UPI00000AA9BA; EMBL:AC004705; NID:g3252804;
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1756 <STO>
A:Cross-references: UNIPARC:UPI00000AA9BA; GB:AE002093; NID:g3252818; PIDN:AAC24188.1; C
C:Genetics:
A:Gene: At2g14770; F26C24.9
A:Map position: 2
A:Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 725/1; 758/3; 778/3; 812/3; 902/2; 9

Query Match 96.4%; Score 27; DB 2; Length 1756;
Best Local Similarity 66.7%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

Qy 1 VRDQIXXXL 9
Db 488 VRDQIRAAAL 496

RESULT 8
T01194
hypothetical protein F21E10.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: T01194
R:Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A:Description: The sequence of A. thaliana F21E10.
A:Reference number: Z14258
A:Accession: T01194
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-211 <DAV>
A:Cross-references: UNIPROT:O65247; UNIPARC:UPI00000A324E; EMBL:AF058914; NID:g3047074;
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F21E10.9
A:Map position: 5

Query Match 92.9%; Score 26; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 42; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 3;

Qy 1 VRDQIXXXL 9
Db 187 VRDQIQITL 195

RESULT 9
AD1062
```

protein kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain C)
 C/Species: Salmonella enterica subsp. enterica serovar Typhi
 A/Note: This species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AD1062
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Paratyphi A
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AD1062
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-357 <PAR>
 A/Cross-references: UNIPARC:UPI000005A993; GB:AL513382; PIDN:CAD06945.1; PTD:gl6505592; C/Genetics:
 A/Gene: STV4823

Query Match 92.9%; Score 26; DB 2; Length 357;
 Best Local Similarity 66.7%; Pred. No. 76;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 255 VRDQIAKYL 263

RESULT 10
 B36329
 hypothetical protein 2 - cabbage looper transposon TED (fragment)
 C/Species: Trichoplusia ni (cabbage looper)
 C/Date: 01-Feb-1991 #sequence_revision 01-Feb-1991 #text_change 30-Sep-1993
 C/Accession: B36329
 R/Friesen, P.D.; Nissen, M.S.
 Mol. Cell. Biol. 10, 3067-3077, 1990
 A/Title: Gene organization and transcription of TED, a lepidopteran retrotransposon integrated into the genome of Trichoplusia ni
 A/Reference number: A36329; MUID:90258898; PMID:1692964
 A/Accession: B36329
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1236 <PRI>
 A/Cross-references: UNIPARC:UPI0000179E4E; GB:M32662

Query Match 92.9%; Score 26; DB 2; Length 1236;
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 306 VRDQITKML 314

RESULT 11
 D85066
 hypothetical protein AT4G05280 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C/Accession: D85066
 R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory
 Nature 402, 769-777, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001; MUID:20083488; PMID:10617198
 A/Accession: D85066
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1312 <STO>
 A/Cross-references: UNIPROT:Q9MOW7; UNIPARC:UPI000000A5C54; GB:NC_001268; NID:G7267288; F723381
 C/Genetics:
 A/Gene: AT4G05280
 A/Map position: 4

Query Match 92.9%; Score 26; DB 2; Length 1312;
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 434 VRDQIRAVL 442

RESULT 12
 S09866
 hypothetical protein UL101 - human cytomegalovirus (strain AD169)
 C/Species: human cytomegalovirus, human herpesvirus 5
 A/Note: host Homo sapiens (man)
 C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C/Accession: S09866
 R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horenell, T.; M.; Barrell, B.G.
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
 A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
 A/Reference number: S09749; MUID:90269039; PMID:2161319
 A/Accession: S09866
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-115 <CHE>
 A/Cross-references: UNIPROT:PI6826; UNIPARC:UPI0000137C06; EMBL:X17403; NID:G59591; PIDN:989
 A/Note: this sequence was submitted to the EMBL Data Library, December 1989
 A/Note: this reading frame extends between two stop codons and does not begin with a start codon
 C/Superfamily: human cytomegalovirus hypothetical protein UL101

Query Match 89.3%; Score 25; DB 2; Length 115;
 Best Local Similarity 55.6%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 73 VRDQLSQL 81

RESULT 13
 D83620
 hypothetical protein PA0201 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C/Accession: D83620
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: D83620
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-193 <STO>
 A/Cross-references: UNIPROT:Q915T7; UNIPARC:UPI000000C4F91; GB:AE004458; GB:AE004091; NID:4091
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA0201

Query Match 89.3%; Score 25; DB 2; Length 193;
 Best Local Similarity 55.6%; Pred. No. 69;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 87 VRDQVRGAL 95

RESULT 14
 T23381
 hypothetical protein K07A1.7 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T23381

R;Percy, C.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19734

A;Accession: T23381

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-243 <WIL>

A;Cross-references: UNIPROT:P09012; UNIPARC:UPI0000082F5A; EMBL:Z81097; PIDN:CAB03168.1;

A;Experimental source: clone K07A1

C;Genetics:

A;Gene: CESP:K07A1.7

A;Map position: 1

A;Introns: 27/3; 113/3; 187/1

Query Match 89.3%; Score 25; DB 2; Length 243;

Best Local Similarity 55.6%; Pred. No. 89;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

||||: |

Db 206 VRDQLATQL 214

RESULT 15

peroxidase (EC 1.11.1.7), neutral - horseradish

C;Species: Armoracia rusticana (horseradish)

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C;Accession: S14268

R;Bartonek-Roxa, E.; Eriksson, H.; Mattiasson, B.

Biochim. Biophys. Acta 1088, 245-250, 1991

A;Title: The cDNA sequence of a neutral horseradish peroxidase.

A;Reference number: S14268; MUID:91159476; PMID:2001399

A;Accession: S14268

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-327 <BAR>

A;Cross-references: UNIPROT:Q42517; UNIPARC:UPI0000131635; GB:X57564; NID:gl6095; PIDN:C

C;Superfamily: peroxidase

C;Keywords: glycoprotein; heme; iron; metalloprotein; oxidoreductase

F;39-116/Disulfide bonds: #status predicted

F;66/Active site: Arg #status predicted

F;70.193/Binding site: heme iron (His) (axial ligands) #status predicted

F;72-77/Disulfide bonds: #status predicted

F;122-323/Disulfide bonds: #status predicted

F;200-232/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 327;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

||||: |

Db 46 VRDQVKIAL 54

RESULT 16

A25228

protein kinase CDC7 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)

N;Alternate names: cell division control protein CDC7; protein D2855; protein YDL017W

C;Species: Saccharomyces cerevisiae

C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 31-Dec-2004

C;Accession: A25228; J70485; S29519; S52502; S67549

R;Patterson, M.; Schafani, R.A.; Fangman, W.L.; Rosamond, J.

Mol. Cell. Biol. 6, 1590-1598, 1986

A;Title: Molecular characterization of cell cycle gene CDC7 from Saccharomyces cerevisiae

A;Reference number: A25228; MUID:87064440; PMID:3537706

A;Accession: A25228

A;Molecule type: DNA

A;Residues: 1-507 <PAT>

A;Cross-references: UNIPROT:P06243; UNIPARC:UPI000012725A; EMBL:M12624; NID:gl71201; PID

R;Bahman, M.; Buck, V.; White, A.; Rosamond, J.

Biochim. Biophys. Acta 951, 335-343, 1988

A;Title: Characterisation of the CDC7 gene product of Saccharomyces cerevisiae as a prot

A;Reference number: J70485; MUID:89088247; PMID:2850010

A;Accession: J70485

A;Molecule type: DNA

A;Residues: 1-138 <BAH>

A;Cross-references: UNIPARC:UPI0000168B5F; EMBL:X14164; NID:g3511; PIDN:CAA32369.1; PID:

R;Ham, J.; Moore, D.; Rosamond, J.; Johnston, I.R.

Nucleic Acids Res. 17, 5781-5792, 1989

A;Title: Transcriptional analysis of the CDC7 protein kinase gene of Saccharomyces cerev

A;Reference number: S29519; MUID:89345168; PMID:2668893

A;Accession: S29519

A;Molecule type: DNA

A;Residues: 1-20 <HAM>

A;Cross-references: UNIPARC:UPI000017A44F; EMBL:X15362

R;Andre, B.; Visgers, S.; Urrestarazu, L.

submitted to the EMBL Data Library, February 1995

A;Description: The sequence of a 42 kb segment located on the left arm of chromosome IV

A;Reference number: S52492

A;Accession: S52502

A;Molecule type: DNA

A;Residues: 1-507 <AND>

A;Cross-references: UNIPARC:UPI000012725A; EMBL:Z48432; NID:g683669; PIDN:CAA88342.1; PI

R;Urrestarazu, L.A.; Andre, B.; Visgers, S.

submitted to the Protein Sequence Database, July 1996

A;Reference number: S67535

A;Accession: S67549

A;Molecule type: DNA

A;Residues: 1-507 <URR>

A;Cross-references: UNIPARC:UPI000012725A; EMBL:Z74065; NID:gl430982; PIDN:CAA98574.1; P

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:CDC7

A;Cross-references: SGD:S0002175; MIPS:YDL017W

A;Map position: 4L

C;Keywords: ATP; cell cycle control; phosphotransferase; serine/threonine-specific prote

F;31-394/Domain: protein kinase homology <KIN>

F;39-47/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 507;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9

||||: |

Db 112 VRDQVIATL 120

RESULT 17

F82192

ABC transporter, ATP-binding protein VC1499 [imported] - Vibrio cholerae (strain N16961

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: F82192

R;Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: F82192

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-610 <HEI>

A;Cross-references: UNIPROT:Q9KRY4; UNIPARC:UPI00000C2FE1; GB:AE004228; GB:AE003852; NI

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1499

A;Map position: 1

Query Match

Best Local Similarity 89.3%; Score 25; DB 2; Length 610;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 444 IRDQIAVVL 452

RESULT 18
AF3016
Na+/H+ antiporter Atu3738 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AF3016
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AF3016
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-621 <KUR>
A/Cross-references: UNIPROT:Q8U9J0; UNIPARC:UPI0000164865; GB:AE008689; PIDN:AAL44548.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu3738
A/Map position: linear chromosome

Query Match 89.3%; Score 25; DB 2; Length 621;
Best Local Similarity 44.4%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 145 IRDQVRSTL 153

RESULT 19
B98268
Probable sodium/hydrogen antiporter PA5021 [imported] - Agrobacterium tumefaciens (strain
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: B98268
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: B98268
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-642 <KUR>
A/Cross-references: UNIPROT:Q8U9J0; UNIPARC:UPI00000022F3; GB:AE007870; PIDN:AAK99668.1;
C/Genetics:
A/Gene: AGR_L_2194
A/Map position: linear chromosome

Query Match 89.3%; Score 25; DB 2; Length 642;
Best Local Similarity 44.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 166 IRDQVRSTL 174

RESULT 20
F89895
ATP-dependent DNA helicase [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: F89895
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: F89895
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-686 <KUR>
A/Cross-references: UNIPROT:Q99UP1; UNIPARC:UPI000000CA9C1; GB:BA000018; PID:gl3701027;
A/Experimental source: strain N315
C/Genetics:
A/Gene: recG
C/Superfamily: DNA helicase recG

Query Match 89.3%; Score 25; DB 2; Length 686;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 166 IRDQIRQAL 174

RESULT 21
T00122
hypothetical 88.9K protein - *Leptospira interrogans*
C/Species: *Leptospira interrogans*
C/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C/Accession: T00122
R/Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998
A/Title: Physical and genetic maps of the *Leptospira interrogans* serovar *icterohaemorrhag*
A/Reference number: Z14115; MUID:98332717; PMID:9666070
A/Accession: T00122
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-780 <TAK>
A/Cross-references: UNIPROT:P97040; UNIPARC:UPI000000BCBC4; EMBL:AB010203; NID:g2780763;
A/Experimental source: strain Ictero No.1; substrain *icterohaemorrhagiae*
C/Superfamily: *Leptospira interrogans* hypothetical 88.9K protein

Query Match 89.3%; Score 25; DB 2; Length 780;
Best Local Similarity 55.6%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 734 IRDQIAEFL 742

RESULT 22
T21910
hypothetical protein F37B12.4 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21910
R/Thomas, K.
submitted to the EMBL Data Library, September 1995
A/Reference number: Z19486
A/Accession: T21910
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1430 <WIL>
A/Cross-references: UNIPROT:O01049; UNIPARC:UPI0000080F77; EMBL:Z54218; PIDN:CAA90957.1;
A/Experimental source: clone F37B12
C/Genetics:
A/Gene: CESP:F37B12.4
A/Map position: 2

A;Introns: 319/2; 666/1; 714/2; 1011/2; 1065/3; 1116/2; 1189/3; 1214/3; 1283/2

Query Match 89.3%; Score 25; DB 2; Length 1430;
Best Local Similarity 55.6%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 135 IRDQIKSL 143

RESULT 23

T22812

hypothetical protein ZC116.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T22812; T27494

R;Burton, J.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z19618

A;Accession: T22812

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-3871 <W1L>

A;Cross-references: UNIPROT:Q20911; UNIPARC:UPI00000821EA; EMBL:Z74473; PIDN:CAA98952.1;

A;Experimental source: clone F56H9

R;Smyle, R.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z20376

A;Accession: T27494

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-3871 <W12>

A;Cross-references: UNIPARC:UPI00000821EA; EMBL:Z74046; PIDN:CAA98557.1; GSPDB:GN000023;

A;Experimental source: clone ZC116

C;Genetics:

A;Gene: CESP:ZC116.3

A;Map position: 5

A;Introns: 29/3; 66/2; 244/1; 359/3; 422/1; 549/1; 572/1; 728/1; 820/1; 889/3; 1189/1; 1

2763/1; 2809/1; 2906/3; 3051/3; 3176/1; 3254/1; 3341/2; 3536/1; 3565/2; 3582/3; 3609/1;

C;Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology

Query Match 89.3%; Score 25; DB 2; Length 3871;

Best Local Similarity 66.7%; Pred. No. 2e+03; 3; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 1714 VRDQIGFVL 1722

RESULT 24

G82442

Su1 family protein VCA0570 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: G82442

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.

l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: G82442

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-103 <HEI>

A;Cross-references: UNIPROT:Q9KMI9; UNIPARC:UPI0000003588; GB:AE004388; GB:AE003853; NID

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VCA0570

A;Map position: 2

C;Superfamily: translation initiation factor SUI

Query Match 85.7%; Score 24; DB 2; Length 103;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

Db 83 VRDQLKTL 91

RESULT 25

A87425

hypothetical protein CC1416 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: A87425

R;Niernan, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

n.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87425

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-109 <STO>

A;Cross-references: UNIPROT:Q9A8D8; UNIPARC:UPI000000C739E; GB:AE005673; NID:gi3422777; P

C;Genetics:

A;Gene: CC1416

Query Match 85.7%; Score 24; DB 2; Length 109;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5

Db 31 VRDQI 35

RESULT 26

T39825

hypothetical protein SPBC19F8.02 - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T39825

R;Beck, A.; Reinhardt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, May 1998

A;Reference number: Z21883

A;Accession: T39825

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-166 <BEC>

A;Cross-references: UNIPROT:O60166; UNIPARC:UPI000006C7AF; EMBL:AL023594; NID:ei293398;

A;Experimental source: strain 972h-; cosmid c19F8

C;Genetics:

A;Gene: SPDB:SPBC19F8.02

A;Map position: 2

Query Match 85.7%; Score 24; DB 2; Length 166;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5

Db 162 VRDQI 166

RESULT 27

E83662

hypothetical protein BH0101 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C;Accession: E83662

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A;Reference number: AB3650; MUID:20512582; PMID:11058132
A;Accession: E83662
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <S70>
A;Cross-references: UNIPROT:Q9JWP8; UNIPARC:UPI000000C492E; GB:AP001507; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0101
C;Superfamily: *Bacillus subtilis* hypothetical protein yach

Query Match 85.7%; Score 24; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|
|
|
|
Db 161 VRDQI 165

RESULT 28
S66113
hypothetical protein yach - *Bacillus subtilis*
C;Species: *Bacillus subtilis*
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66113; A69741
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA Res. 1, 1-14, 1994
A;Title: Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chrom
A;Reference number: S65967; MUID:96051385; PMID:7584024
A;Accession: S66113
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-185 <OGA>
A;Cross-references: UNIPROT:P37569; UNIPARC:UPI000005PE01; EMBL:D26185; NID:G467326; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:9804033; PMID:9384377
A;Accession: A69741
A;Status: nucleic acid sequence not shown; translation not shown
A;Residues: 1-185 <KUN>
A;Molecule type: DNA
A;Cross-references: UNIPARC:UPI000005PE01; GB:Z99104; GB:AL009126; NID:G2632267; PIDN:CA
A;Experimental source: strain 168
C;Genetics:
A;Gene: yach
A;Start codon: TTG
C;Superfamily: *Bacillus subtilis* hypothetical protein yach

Query Match 85.7%; Score 24; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|
|
|
|
Db 164 VRDQI 168

RESULT 29

S39401

hydrogenase maturation factor hupD [similarity] - *Bradyrhizobium japonicum*C;Species: *Bradyrhizobium japonicum*

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: S39401

R;van Soom, C.; Browaeys, J.; Verreth, C.; Vanderleyden, J.

J. Mol. Biol. 234, 508-512, 1993

A;Title: Nucleotide sequence analysis of four genes, hupC, hupD, hupF and hupG, downstre

A;Reference number: S39400; MUID:94047099; PMID:8230232

A;Accession: S39401

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-193 <VAN>

A;Cross-references: UNIPROT:Q45251; UNIPARC:UPI000016E70D; EMBL:Z21948; NID:G311336; PID

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993

C;Superfamily: [NiFe]-hydrogenase maturation protease

Query Match 85.7%; Score 24; DB 1; Length 193;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5

|
|
|
|

Db 142 VRDQI 146

RESULT 30

B82601

transcription regulator AcrR family XF2085 [imported] - *Xylella fastidiosa* (strain 9a5c)C;Species: *Xylella fastidiosa*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A;Reference number: A8215; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: B82601

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-205 <STM>

A;Cross-references: UNIPROT:Q9PBQ5; UNIPARC:UPI000000C2924; GB:AE004024; GB:AE003849; NID

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H

as-Neto, E.; Docena, C.; El-Dorfi, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2085

Query Match 85.7%; Score 24; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5

|
|
|
|

Db 19 VRDQI 23

RESULT 31

F81658
 type III secretion translocase sctL TC0850 [imported] - Chlamydia muridarum (strain Nigg)
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: F81658
 R:Read, T.D.; Brunham, R.C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A:Reference number: AB1500; MUID:20150255; PMID:10684935
 A:Accession: F81658
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-235 <TE>
 A:Cross-references: UNIPROT:Q9R9H9; UNIPARC:UPI0000057AA8; GB:AE002160; NID
 A:Experimental source: strain Nigg (MoPn)
 C:Genetics:
 A:Gene: TC0850

Query Match 85.7%; Score 24; DB 2; Length 235;
 Best Local Similarity 55.6%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 :|||||
 DB 98 MRDQIKSSL 106

RESULT 32
 E70104
 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) (plsC) - Lyme disease spirochete
 N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acetyltransferase
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 16-Aug-2004
 C:Accession: E70104; I40293
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Accession: E70104
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-250 <KLE>
 A:Cross-references: UNIPROT:Q59188; UNIPARC:UPI0000057312; GB:AE001117; GB:AE000783; NID
 A:Experimental source: strain B31
 R:Ojaimi, C.; Davidson, B.E.; Saint Girons, I.; Old, I.G.
 Microbiology 140, 2931-2940, 1994
 A:Title: Conservation of gene arrangement and an unusual organization of rRNA genes in B
 A:Reference number: I40241; MUID:95111614; PMID:7812434
 A:Accession: I40293
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 116-250 <RES>
 A:Cross-references: UNIPARC:UPI000016E69A; GB:L32861; NID:G520779; PIDN:AAC41407.1; PID:
 C:Genetics:
 A:Gene: plsC
 C:Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match 85.7%; Score 24; DB 2; Length 250;
 Best Local Similarity 55.6%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
 :|||||
 DB 237 IRDQIVKKL 245

RESULT 33
 AE0127

Conserved hypothetical protein YP01038 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AE0127
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AE0127
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <KUR>
 A:Cross-references: UNIPROT:Q8ZH72; UNIPARC:UPI00000DCCEP; GB:AL590842; PIDN:CAC89880.1;
 C:Genetics:
 A:Gene: YP01038
 C:Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 257;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
 :||||
 DB 39 VRDQI 43

RESULT 34
 A69965
 ketoacyl reductase homolog yqjQ - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
 C:Accession: A69965
 R:Kunst, P.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bercei C.; Bron, S.; Brouillet, S.; Pruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleri ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinola, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Mauei Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, Akeuchi, M.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis. A:Title: The complete genome sequence of the Gram-positive bacterium A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-259 <KUN>
 A:Cross-references: UNIPROT:P54554; UNIPARC:UPI00000606B5; GB:Z99116; GB:AL0009126; NID:
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yqjQ
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology F;7-183/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 85.7%; Score 24; DB 2; Length 259;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRDQI 5
 :||||
 DB 74 VRDQI 78

RESULT 35
 AG0861

probable RNA pseudouridylylate synthase yqcB [imported] - Salmonella enterica subsp. ente

C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: this species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0861
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <PAR>
A;Cross-references: UNIPARC:UPI000005A3BD; GB:AL513382; PIDN:CAD06078.1; PID:g16504044;
C;Genetics:
A;Gene: YQCB
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|||||
Db 39 VRDQI 43

RESULT 36
C91085
hypothetical protein ECs3651 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C91085
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gastwara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C91085
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <HAY>
A;Cross-references: UNIPROT:Q8X6T6; UNIPARC:UPI000000DB23; GB:BA0000007; PIDN:BA037074.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs3651
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|||||
Db 39 VRDQI 43

RESULT 37
E85930
hypothetical protein 24107 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL9
C;Species: *Escherichia coli*
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85930
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:11206551
A;Accession: E85930
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-260 <STO>
A;Cross-references: UNIPROT:Q8X6T6; UNIPARC:UPI000000DOB23; GB:AE005174; NID:g12517267; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4107
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|||||
Db 39 VRDQI 43

RESULT 38
C65061
hypothetical protein b2791 - *Escherichia coli* (strain K-12)
C;Species: *Escherichia coli*
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C65061
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of *Escherichia coli* K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C65061
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-260 <BLAT>
A;Cross-references: UNIPROT:Q46918; UNIPARC:UPI000013BED9; GB:AE000363; GB:U000096; NID:G
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: conserved hypothetical protein HI0176

Query Match 85.7%; Score 24; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|||||
Db 39 VRDQI 43

RESULT 39
T22929
hypothetical protein F58G1.5 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22929
R;Smyle, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19639
A;Accession: T22929
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-302 <WIL>
A;Cross-references: UNIPROT:O62271; UNIPARC:UPI00000801FB; EMBL:Z81556; PIDN:CAB04520.1;
A;Experimental source: clone F58G1
C;Genetics:
A;Gene: CEGP:F58G1.5
A;Map position: 2
A;Introns: 89/2; 127/3; 162/1; 231/3; 262/3
C;Superfamily: *Caenorhabditis elegans* hypothetical protein C09E8.2

Query Match 85.7%; Score 24; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
|||||
Db 37 VRDQI 41

RESULT 40

CPCVB
C:Species: Astacus astacus, Astacus fluviatilis (broad-fingered crayfish)
C:Date: 30-Jun-1998 #sequence_revision 30-Jun-1998 #text_change 09-Jul-2004
C:Accession: A05141; A14178
R:Ticani, K.; Ericsen, L.H.; Kumar, S.; Jakob, F.; Neurath, H.; Zwilling, R.
Biochemistry 23, 1245-1250, 1984
A:Title: Amino acid sequence of crayfish (Astacus fluviatilis) carboxypeptidase B.
A:Reference number: A90482
A:Accession: A05141
A:Molecule type: protein
A:Residues: 1-303 <TIT>
A:Cross-references: UNIPROT:P04069; UNIPARC:UPI00001271CA
C:Superfamily: carboxypeptidase
C:Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein digestion; zinc
F:241,264/Active site: Tyr, Glu #status predicted

Query Match 85.7%; Score 24; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 VRDQI 5
|||
DB 171 VRDQI 175

RESULT 41

H87447
conserved hypothetical protein CC1601 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87447
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <STO>
A:Cross-references: UNIPROT:Q9A7W8; UNIPARC:UPI00000C7439; GB:AE005673; NID:g13422598; E
C:Genetics:
A:Gene: CC1601

Query Match 85.7%; Score 24; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 VRDQI 5
|||
DB 94 VRDQI 98

RESULT 42

H87271
general secretion pathway protein K [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87271
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87271
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <STO>

A:Cross-references: UNIPROT:Q9ABP6; UNIPARC:UPI00000C6F67; GB:AE005673; NID:g13421299; P
C:Genetics:
A:Gene: CC0180
C:Superfamily: outK protein

Query Match 85.7%; Score 24; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 VRDQI 5
|||
DB 275 VRDQI 279

RESULT 43

T29369
hypothetical protein ZC404.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29369
R:Bentley, D.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid ZC404.
A:Reference number: Z20614
A:Accession: T29369
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-351 <BEN>
A:Cross-references: UNIPROT:Q23287; UNIPARC:UPI0000001CFD; EMBL:U55363; PIDN:AAA97963.1;
A:Experimental source: strain Bristol N2; clone ZC404
C:Genetics:
A:Gene: CESP:ZC404.8
A:Map position: 5
A:Introns: 17/2; 52/2; 73/2; 312/1

Query Match 85.7%; Score 24; DB 2; Length 351;
Best Local Similarity 55.6%; Pred. No. 2.4e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 5; Conservative 1

QY 1 VRDQIXXXL 9
:|||
DB 267 LRDQITALL 275

RESULT 44

T04917
hypothetical protein T10114.190 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04917
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15389
A:Accession: T04917
A:Molecule type: DNA
A:Residues: 1-369 <BEV>
A:Cross-references: UNIPROT:O49641; UNIPARC:UPI00000A4B3A; EMBL:AL021712
A:Experimental source: cultivar Columbia; BAC clone T10114
C:Genetics:
A:Map position: 4
A:Introns: 183/2; 220/3; 278/1
A:Note: T10114.190

Query Match 85.7%; Score 24; DB 2; Length 369;
Best Local Similarity 55.6%; Pred. No. 2.6e+02; Mismatches 3; Indels 0; Gaps 0;
Matches 5; Conservative 1

QY 1 VRDQIXXXL 9
:|||
DB 48 IRDQINILL 56

RESULT 45

S45107
 hypothetical protein 1 - Erwinia carotovora
 C/Species: Erwinia carotovora
 C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C/Accession: S45107
 R/Golby, P.; Jones, S.E.; Stephens, S.; Reeves, P.J.; Bycroft, B.; Stewart, G.; Williams submitted to the EMBL Data Library, May 1994
 A/Description: Global regulation of Erwinia carotovora exoenzyme virulence factors: mult
 A/Reference number: S45107
 A/Accession: S45107
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-376 <GO>
 A/Cross-references: UNIPROT:Q47417; UNIPARC:UPI000013BEDA; EMBL:X79474; NID:g496597; PID

Query Match 85.7%; Score 24; DB 2; Length 376;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
 Db 153 VRDQI 157
 |||||

RESULT 46
 T32710
 hypothetical protein T22D1.10 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T32710
 R/Geisel, C.; Bradshaw, H.; Hawkins, M.
 submitted to the EMBL Data Library, December 1997
 A/Description: The sequence of C. elegans cosmid T22D1.
 A/Reference number: T21211
 A/Accession: T32710
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-448 <GEI>
 A/Cross-references: UNIPROT:Q9GZH2; UNIPARC:UPI000007D4F4; EMBL:AF039052; PIDN:AA94278.
 A/Experimental source: strain Bristol N2; clone T22D1
 C/Genetics:
 A/Gene: CESP:T22D1.10
 A/Map position: 4
 A/Introns: 14/3; 60/3; 103/3; 196/1; 326/2; 377/2
 C/Superfamily: conserved hypothetical protein YDR190c

Query Match 85.7%; Score 24; DB 2; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
 Db 268 VRDQI 272
 |||||

RESULT 47
 S61029
 hypothetical protein YPL235w - yeast (Saccharomyces cerevisiae)
 N/Alternate names: hypothetical protein PI060
 C/Species: Saccharomyces cerevisiae
 C/Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
 C/Accession: S61029; S61699; S65260; S65264
 R/Pohl, T.M.
 submitted to the EMBL Data Library, November 1995
 A/Reference number: S61010
 A/Accession: S61029
 A/Molecule type: DNA
 A/Residues: 1-471 <POH>
 A/Cross-references: UNIPROT:Q12464; UNIPARC:UPI0000052E06; EMBL:Z67751; NID:g1061234; PI
 R/Urrastazu, L.A.
 submitted to the EMBL Data Library, December 1995
 A/Reference number: S61699
 A/Accession: S61699

A/Molecule type: DNA
 A/Residues: 1-471 <URR>
 A/Cross-references: UNIPARC:UPI0000052E06; EMBL:X94561; NID:g1181252; PIDN:CAA64252.1; P
 R/Urrastazu, L.A.; Vissers, S.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S65251
 A/Accession: S65260
 A/Molecule type: DNA
 A/Residues: 1-471 <URF>
 A/Cross-references: UNIPARC:UPI0000052E06; EMBL:Z73591; NID:g1370485; PIDN:CAA97952.1; P
 A/Experimental source: strain S288C (AB972)
 R/Pohl, T.M.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64899
 A/Accession: S65264
 A/Molecule type: DNA
 A/Residues: 1-471 <POW>
 A/Cross-references: UNIPARC:UPI0000052E06; EMBL:Z73591; NID:g1370485; PIDN:CAA97952.1; P
 A/Experimental source: strain S288C (AB972)
 C/Genetics:
 A/Gene: SGD:RVB2
 A/Cross-references: SGD:S0006156
 A/Map position: 16L
 C/Superfamily: conserved hypothetical protein YDR190c

Query Match 85.7%; Score 24; DB 2; Length 471;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRDQI 5
 Db 269 VRDQI 273
 |||||

RESULT 48
 S64837
 hypothetical protein YLR015w - yeast (Saccharomyces cerevisiae)
 N/Alternate names: hypothetical protein LI583
 C/Species: Saccharomyces cerevisiae
 C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
 C/Accession: S64837; S59274; S48527
 R/Vandenbol, M.; Portetelle, D.; Hilger, F.
 submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64742
 A/Accession: S64837
 A/Molecule type: DNA
 A/Residues: 1-505 <VAN>
 A/Cross-references: UNIPROT:P43132; UNIPARC:UPI0000052PF0; EMBL:Z73187; NID:g1360313; PI
 A/Experimental source: strain S288C
 R/Saville, S.P.; Atkinson, S.; Jamieson, L.; Pocklington, M.J.; Orr, E.
 submitted to the EMBL Data Library, August 1995
 A/Description: A 7.8kb fragment from chromosome XII of Saccharomyces cerevisiae does not
 A/Reference number: S59270
 A/Accession: S59274
 A/Molecule type: DNA
 A/Residues: 1-464, 'ISLL', 468 <SAV>
 A/Cross-references: UNIPARC:UPI000006AFA4; EMBL:X90564; NID:g975221; PIDN:CAA62158.1; PI
 A/Experimental source: strain S288C
 R/Levin, D.E.; Stevenson, W.D.
 submitted to the EMBL Data Library, July 1994
 A/Description: The S.cerevisiae PKC2 does not exist in the yeast genome.
 A/Reference number: S48527
 A/Accession: S48527
 A/Molecule type: DNA
 A/Residues: 1-472 <LEV>
 A/Cross-references: UNIPARC:UPI0000179EB1; EMBL:L34405
 C/Genetics:
 A/Gene: SGD:BRE2
 A/Cross-references: SGD:S0004005
 A/Map position: 12R

Query Match 85.7%; Score 24; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Search completed: May 12, 2006, 10:52:38
Job time : 19.0231 secs

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	309	VRDQI 313							
RESULT 49									
PQ0646									
ferredoxin-nitrite reductase (EC 1.7.7.1) - <i>Synechococcus</i> sp. (strain PCC 7942)									
C:Species: <i>Synechococcus</i> sp.									
A:Variety: PCC 7942									
C>Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 31-Dec-2004									
C:Accession: S33530; PQ0646; S32143									
R:Luque, I.; Flores, E.; Herrero, A.									
Plant Mol. Biol. 21, 1201-1205, 1993									
A:Title: Nitrite reductase gene from <i>Synechococcus</i> sp. PCC 7942: homology between cyanob									
A:Reference number: S33530; MUID:93257637; PMID:9490140									
A:Accession: S33530									
A:Molecule type: DNA									
A:Residues: 1-512 <LUQ>									
A:Cross-references: UNIPROT:P39661; UNIPARC:UPI00001301CB; EMBL:X67680; NID:g288053; PID									
R:Omata, T.									
Plant Cell Physiol. 32, 151-157, 1991									
A:Title: Cloning and characterization of the nrtA gene that encodes a 45-kDa protein inv									
A:Reference number: PQ0646									
A:Accession: PQ0646									
A:Molecule type: DNA									
A:Residues: 468-512 <OWA>									
A:Cross-references: UNIPARC:UPI0000172138; DDBJ:D00677									
C:Genetics:									
A:Gene: nrt									
C:Superfamily: sulfite/ferredoxin-nitrite reductase									
C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein; oxidoreductase									
F:396,402,437,441/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted									
Query Match 85.7%; Score 24; DB 1; Length 512;									
Best Local Similarity 100.0%; Pred. No. 3.7e+02; Mismatches 0; Indels 0; Gaps 0;									
Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	26	VRDQI 30							
RESULT 50									
G83558									
hypothetical protein PA0696 [imported] - <i>Pseudomonas aeruginosa</i> (strain PA01)									
C:Species: <i>Pseudomonas aeruginosa</i>									
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004									
C:Accession: G83558									
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B									
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,									
.; Lory, S.; Olson, M.V.									
Nature 406, 959-964, 2000									
A:Title: Complete genome sequence of <i>Pseudomonas aeruginosa</i> PA01, an opportunistic patho									
A:Reference number: A82950; MUID:20437337; PMID:10984043									
A:Accession: G83558									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-568 <STO>									
A:Cross-references: UNIPROT:Q915N0; UNIPARC:UPI00000C5107; GB:AE004505; GB:AE004091; NID									
A:Experimental source: strain PA01									
C:Genetics:									
A:Gene: PA0696									
Query Match 85.7%; Score 24; DB 2; Length 568;									
Best Local Similarity 100.0%; Pred. No. 4.2e+02; Mismatches 0; Indels 0; Gaps 0;									
Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	VRDQI 5							
Db	96	VRDQI 100							

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 63.6923 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-33
Perfect score: 28
Sequence: 1 VRDQIXXXL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt 05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	232	Q6MCX7_BDRBA	Q6mcx7 bdellovibri
2	27	96.4	455	Q7QAB5_ANOGA	Q7qab5 anopheles g
3	27	96.4	543	Q9ZBH5_ATH	Q9zbh5 arabidopsis
4	27	96.4	1113	Q9ML80_ATH	Q9ml80 arabidopsis
5	27	96.4	1119	Q9LK04_ATH	Q9lk04 arabidopsis
6	27	96.4	1132	Q9S926_ATH	Q9s926 arabidopsis
7	27	96.4	1198	Q9C6N9_ATH	Q9c6n9 arabidopsis
8	27	96.4	1201	Q9C607_ATH	Q9c607 arabidopsis
9	27	96.4	1305	Q9ZU21_ATH	Q9zu21 arabidopsis
10	27	96.4	1444	Q9LTC8_ATH	Q9ltc8 arabidopsis
11	27	96.4	1444	Q9SFV0_ATH	Q9sfv0 arabidopsis
12	27	96.4	1472	Q9SL11_ATH	Q9sl11 arabidopsis
13	27	96.4	1756	Q80975_ATH	Q80975 arabidopsis
14	26	92.9	211	Q65247_ATH	Q65247 arabidopsis
15	26	92.9	218	Q7XJK7_ATH	Q7xjk7 arabidopsis
16	26	92.9	272	Q5WH83_BACSK	Q5wh83 bacillus cl
17	26	92.9	334	Q7NNK0_GLOVI	Q7nnk0 gloeobacter
18	26	92.9	357	Q8Z110_SALTI	Q8z110 salmonella
19	26	92.9	486	Q8MYV0_DROME	Q8myv0 drosophila
20	26	92.9	486	Q9VM68_DROME	Q9vm68 drosophila
21	26	92.9	516	Q4PCCI_USTMA	Q4pccl ustilago ma
22	26	92.9	594	Q4Q7Y6_LEIMA	Q4q7y6 leishmania
23	26	92.9	642	Q89S89_BRAJA	Q89s89 bradyrhizob
24	26	92.9	743	Q5LX00_SILPO	Q5lx00 silicibacte
25	26	92.9	896	Q7UYC9_RHOBA	Q7uyc9 rhodopirell
26	26	92.9	1084	Q65353_NPVAC	Q65353 autographa
27	26	92.9	1105	Q4P3H5_USTMA	Q4p3h5 ustilago ma
28	26	92.9	1191	Q411R0_GIBZE	Q411r0 gibberella
29	26	92.9	1312	Q9M0W7_ATH	Q9m0w7 arabidopsis
30	26	92.9	1407	Q61Z65_CAEBR	Q61z65 caenorhabdi
31	26	92.9	2240	Q51WD5_MAGGR	Q51wd5 magnaporthe

32	25	89.3	2	Q76VV2_9CAUD	Q76vv2 bacterioph
33	25	89.3	2	Q6RXD6_HCMV	Q6rxd6 human cytom
34	25	89.3	115	UL101_HCMVA	UL101 human cytom
35	25	89.3	130	Q58J60_9ACTO	Q58j60 streptomyce
36	25	89.3	139	Q4XM22_PLACH	Q4xm22 plasmodium
37	25	89.3	193	Q916T7_PSEAE	Q916t7 pseudomonas
38	25	89.3	225	Q65L06_BACLD	Q65l06 bacillus li
39	25	89.3	229	Q6VT06_VIRU	Q6vt06 vibrio para
40	25	89.3	243	Q90912_CAEBL	Q90912 caenorhabdi
41	25	89.3	250	EVG1_DROME	Q9vss7 drosophila
42	25	89.3	250	Q8MZ83_DROME	Q8mz83 drosophila
43	25	89.3	291	Q72MR9_LEPIC	Q72mr9 leptospira
44	25	89.3	291	Q8EZE2_LEPIN	Q8eze2 leptospira
45	25	89.3	296	Q8G822_BIFLO	Q8g822 bifidobacte
46	25	89.3	327	PERN_ARMKU	Q42517 armoracia r
47	25	89.3	358	Q61Y40_CAEBR	Q61y40 caenorhabdi
48	25	89.3	422	Q4WFK8_ASPFU	Q4wfk8 aspergillus
49	25	89.3	437	Q8DMW3_VIBVU	Q8dmw3 vibrio vuln
50	25	89.3	437	Q7MLK9_VIBVU	Q7mlk9 vibrio vuln
51	25	89.3	452	Q9LJ56_ATH	Q9lj56 arabidopsis
52	25	89.3	495	Q6D280_ERWCT	Q6d280 erwinia car
53	25	89.3	504	Q6FN90_CANGA	Q6fn90 candida gla
54	25	89.3	507	CDCT_YEAST	Q6q280 arwinda car
55	25	89.3	516	Q5VEI8_SOLTU	P06243 saccharomyc
56	25	89.3	522	Q525T3_MAGGR	Q525t3 magnaporthe
57	25	89.3	530	GUAA_BORPA	Q7wkl3 bordetella
58	25	89.3	530	GUAA_BORPE	Q7wv46 bordetella
59	25	89.3	530	GUAA_BORPA	Q7vwm4 bordetella
60	25	89.3	551	G6PI_SYNPX	Q8pyb9 methanosarc
61	25	89.3	587	Q8PYB9_METMA	Q7u6t0 synecococc
62	25	89.3	610	Q9KRY4_VIBCH	Q8kry4 vibrio chol
63	25	89.3	619	Q7DLA2_LEPIN	Q7dl22 leptospira
64	25	89.3	621	Q8U9J0_AGRTS	Q8u9j0 agrobacteri
65	25	89.3	624	Q5WVJ4_LRGPL	Q5wvj4 legionella
66	25	89.3	624	Q5X455_LRGPA	Q5x455 legionella
67	25	89.3	624	Q5ZUE3_LSGPH	Q5zue3 legionella
68	25	89.3	635	Q4H812_9DEIO	Q4h812 deinococcus
69	25	89.3	642	Q4YN59_PLABE	Q4yn59 plasmodium
70	25	89.3	642	Q7CT69_AGRTS	Q7ct69 agrobacteri
71	25	89.3	686	RECG_STAAN	P64324 staphylococ
72	25	89.3	686	RECG_STAAN	O50581 staphylococ
73	25	89.3	686	RECG_STAAN	O50581 staphylococ
74	25	89.3	686	RECG_STAAN	Q6nx11 staphylococ
75	25	89.3	686	Q6G9Y6_STAAS	Q6g9y6 staphylococ
76	25	89.3	686	Q6GHK8_STAAC	Q6ghk8 staphylococ
77	25	89.3	686	Q5HGK6_STAAC	Q5hgk6 staphylococ
78	25	89.3	737	Q5H511_XANOR	Q5h511 xanthomonas
79	25	89.3	737	Q8PQ22_XANAC	Q8pq22 xanthomonas
80	25	89.3	747	Q4UZD4_XANCP	Q4uzd4 xanthomonas
81	25	89.3	747	Q8PD49_XANCP	Q8pd49 xanthomonas
82	25	89.3	771	Q72AX1_DESVH	Q72ax1 desulfovibr
83	25	89.3	780	Q72VZ2_LEPIC	P97040 leptospira
84	25	89.3	784	Q72VZ2_LEPIC	Q72vz2 leptospira
85	25	89.3	793	Q5AS07_EMENI	Q5as07 leptospira
86	25	89.3	827	Q7UM67_RHOBA	Q7um67 rhodopirell
87	25	89.3	836	Q4FEY1_9VIRU	Q4fey1 cotesia plu
88	25	89.3	853	Q7RDQ3_PLAYO	Q7rdq3 plasmodium
89	25	89.3	960	Q7S6I2_NEUCR	Q7s6i2 neurospora
90	25	89.3	965	Q6SLD5_COCHC	Q6sl55 coenobolus
91	25	89.3	988	P93661_CHEAL	P93661 cheenopodiu
92	25	89.3	1073	Q82QD5_STRPV	Q82qd5 streptomyce
93	25	89.3	1073	Q5CUN7_CRYVP	Q5cun7 cryptospori
94	25	89.3	1779	Q5CPD0_CRYHO	Q5cpd0 cryptospori
95	25	89.3	1880	Q6CAJ0_YARLI	Q6caj0 yarrowia li
96	25	89.3	2183	P90865_CABEL	P90865 caenorhabdi
97	25	89.3	2923	Q7JLZ7_CABEL	Q7jlz7 caenorhabdi
98	25	89.3	3251	Q5K8V2_CRYNE	Q5k8v2 cryptococcu
99	25	89.3	3342	Q55NG7_CRYNE	Q55ng7 cryptococcu
100	25	89.3	3871	Q20911_CABEL	Q20911 caenorhabdi
101	25	89.3	26	Q6V7G4_VIBCH	Q6v7g4 vibrio chol
102	24	85.7	39	Q4TIC1_TETNG	Q4tic1 tetraodon n
103	24	85.7	50	Q4SUG6_TETNG	Q4sug6 tetraodon n
104	24	85.7	50	Q4SUG6_TETNG	Q4sug6 tetraodon n

105	24	85.7	70	2	Q85309_9CAUD	Q853q9_mycobacteri	178	24	85.7	259	1	YQJQ_BACSU	P54554 bacillus su
106	24	85.7	74	2	Q58L62_BRACM	Q58l62_brassica ca	179	24	85.7	259	2	Q4ZT89_PSESY	Q4zt89_pseudomonas
107	24	85.7	84	2	Q58D66_STRP3	Q58d66_streptococc	180	24	85.7	260	1	TRUC_ECOS7	Q8x6t6_escherichia
108	24	85.7	84	2	Q5XD61_STRP6	Q5xd61_streptococc	181	24	85.7	260	1	TRUC_ECOL6	Q8x6t6_escherichia
109	24	85.7	84	2	Q9A140_STRP3	Q9a140_streptococc	182	24	85.7	260	1	TRUC_ECOLI	Q46918_escherichia
110	24	85.7	84	2	Q8P264_STRP8	Q8p264_streptococc	183	24	85.7	260	1	TRUC_SALTI	Q8z439_salmonella
111	24	85.7	88	2	Q95I05_BOVIN	Q95iq5_bos taurus	184	24	85.7	260	1	TRUC_SALTY	Q8zmd5_salmonella
112	24	85.7	95	2	Q4KGS5_PSEFS	Q4kgs5_pseudomonas	185	24	85.7	260	2	Q57KF2_SALCH	Q57kd5_salmonella
113	24	85.7	103	2	Q9KM19_VIBCH	Q9km19_vibrio chol	186	24	85.7	260	2	Q5PEK3_SALPA	Q5pek3_salmonella
114	24	85.7	104	2	Q6N279_RHOPA	Q6n279_rhodopseudo	187	24	85.7	261	2	Q6D8E9_ERWCT	Q6d8e9_erwinia car
115	24	85.7	109	2	Q9A8D8_CAUCR	Q9a8d8_caulobacter	188	24	85.7	262	2	Q5CU68_CRYPV	Q5cu68_cryptospori
116	24	85.7	129	2	Q8NTP3_CORGL	Q8ntp3_corynebacte	189	24	85.7	269	2	Q8CFH8_MOUSE	Q8cfh8_mus musculu
117	24	85.7	130	2	Q9LMS4_ARATH	Q9lms4_arabidopsais	190	24	85.7	270	2	Q9LM35_ARATH	Q9lm35_arabidopsais
118	24	85.7	139	2	Q7V411_PROHM	Q7v411_prochloroco	191	24	85.7	277	2	Q85S29_CRYNV	Q85s29_cryptococcu
119	24	85.7	139	2	Q7U457_SYNXP	Q7u457_synechococc	192	24	85.7	280	2	Q6L1R33_CAEBR	Q6l1r33_caenorbhadi
120	24	85.7	141	2	Q890G2_LACPL	Q890g2_lactobacilli	193	24	85.7	285	2	Q4QPJ5_HAE18	Q4qjp5_haemophilus
121	24	85.7	142	2	Q86DC5_CABEL	Q86dc5_caenorbhadi	194	24	85.7	287	2	Q89L16_BRAJA	Q89l16_bradyrhizob
122	24	85.7	151	2	Q89QJ0_STRCO	Q89qj0_streptomyce	195	24	85.7	292	2	Q6M227_CORGL	Q6m227_corynebacte
123	24	85.7	152	2	Q64G27_CLOIN	Q64g27_clostridium	196	24	85.7	294	2	Q89HJ2_BRAJA	Q89hj2_bradyrhizob
124	24	85.7	157	2	Q6AM17_DESPS	Q6am17_desulfotale	197	24	85.7	295	2	Q4N167_THEPA	Q4n167_theileria p
125	24	85.7	166	2	Q6O166_SCHPO	Q6o166_schizosacch	198	24	85.7	295	2	Q5FM19_LACAC	Q5fm19_lactobacilli
126	24	85.7	168	2	Q5Y4A6_CARPA	Q5y4a6_carica papa	199	24	85.7	296	2	Q74KG0_LACJO	Q74kg0_lactobacilli
127	24	85.7	175	2	Q86HU6_DICDI	Q86hu6_dictyosteli	200	24	85.7	302	2	Q62271_CABEL	Q62271_caenorbhadi
128	24	85.7	175	2	Q8BI02_MOUSE	Q8bi02_mus musculu	201	24	85.7	302	2	Q87H02_VIBPA	Q87h02_vibrio para
129	24	85.7	177	2	Q9JWP8_BACHD	Q9jwp8_bacillus ha	202	24	85.7	303	1	CBPB_ASTFL	P04069_astacus flu
130	24	85.7	182	2	Q5L438_GEOKA	Q5l438_geobacillus	203	24	85.7	306	2	Q8IN95_DROME	Q8in95_drosophila
131	24	85.7	183	2	Q7SC94_NEUCR	Q7sc94_neurospora	204	24	85.7	307	2	Q75H63_ORYSA	Q75h63_oryza sativ
132	24	85.7	185	1	YACH_BACSU	P37569_bacillus su	205	24	85.7	315	2	Q9A7M8_CAUCR	Q9a7m8_caulobacter
133	24	85.7	192	1	HUPD_BRAJA	Q45251_bradyrhizob	206	24	85.7	316	1	CCSA_ADICA	Q85fh6_adiantum ca
134	24	85.7	204	2	Q88HQ3_PSEPK	Q88hq3_pseudomonas	207	24	85.7	316	2	Q6FAS4_ACIAD	Q6fas4_acinetobact
135	24	85.7	204	2	Q8XU11_RALSO	Q8xui1_raistonia s	208	24	85.7	316	2	Q9EQ05_MOUSE	Q9eq05_mus musculu
136	24	85.7	205	2	Q4LYE5_9BURK	Q4lye5_burkholderi	209	24	85.7	317	1	T2N1_NOCAR	P50187_nocardia ae
137	24	85.7	205	2	Q9PB05_XYLFA	Q9pb05_xylella fas	210	24	85.7	319	2	Q4LUR9_9BURK	Q4lur9_burkholderi
138	24	85.7	205	2	Q8PQ91_XANAC	Q8pq91_xanthomonas	211	24	85.7	324	2	Q9ABP6_CAUCR	Q9abp6_caulobacter
139	24	85.7	206	2	Q7Q261_ANOGA	Q7q261_anopheles g	212	24	85.7	325	2	Q50WB0_ENTHI	Q50wb0_entamoeba h
140	24	85.7	207	2	Q4UZK6_XANCP	Q4uzk6_xanthomonas	213	24	85.7	328	2	Q8KPM8_SACER	Q8kpm8_saccharopol
141	24	85.7	207	2	Q8PDC0_XANCP	Q8pdc0_xanthomonas	214	24	85.7	333	2	Q8GM82_ARATH	Q8gm82_arabidopsais
142	24	85.7	209	2	Q9XLM6_9HOMO	Q9xlm6_coniophora	215	24	85.7	338	1	RFC3_ARXAD	Q74111_arxula aden
143	24	85.7	211	2	Q5N3V5_SYNPE	Q5n3v5_synechococc	216	24	85.7	338	2	Q5LNQ8_SILPO	Q5lnq8_silicibacte
144	24	85.7	212	2	Q9XLM0_9HOMO	Q9xlm0_paxillus fi	217	24	85.7	340	2	Q4H9K5_9DEIO	Q4h9k5_deinococcus
145	24	85.7	212	2	Q9XLM1_PAXIN	Q9xlm1_paxillus in	218	24	85.7	342	1	PLSX_LEGPA	Q5xvph3_legionella
146	24	85.7	212	2	Q9XLM7_9HOMO	Q9xlm7_phaeoglycopo	219	24	85.7	342	1	PLSX_LEGPH	Q5xvph6_legionella
147	24	85.7	212	2	Q9XLM8_9HOMO	Q9xlm8_boletinellu	220	24	85.7	342	1	PLSX_LEGPL	Q5wvvc_legionella
148	24	85.7	212	2	Q9XLM9_9HOMO	Q9xlm9_gyrodont liv	221	24	85.7	344	1	SYPA_OCRH	Q8eph4_oceanobacill
149	24	85.7	212	2	Q9XLM1_9HOMO	Q9xlm1_scleroderma	222	24	85.7	345	2	Q5WLF7_BACSK	Q5wlf7_bacillus cl
150	24	85.7	216	2	Q79606_BOLED	Q79606_boletus edu	223	24	85.7	347	2	Q4N2G9_THEPA	Q4n2g9_theileria p
151	24	85.7	219	2	Q9XLM9_9HOMO	Q9xlm9_chalciporus	224	24	85.7	347	2	Q4U9T3_THEAN	Q4u9t3_theileria a
152	24	85.7	221	2	Q79607_9HOMO	Q79607_boletus pal	225	24	85.7	351	2	Q23287_CABEL	Q23287_caenorbhadi
153	24	85.7	221	2	Q79608_9HOMO	Q79608_xerocomus c	226	24	85.7	351	2	Q7UNN0_RHOBA	Q7unn0_rhodopirell
154	24	85.7	221	2	Q9XLM7_9HOMO	Q9xlm7_phylloporus	227	24	85.7	358	2	Q6C137_YARLI	Q6c137_yarrowia li
155	24	85.7	221	2	Q9XLM8_9HOMO	Q9xlm8_chamonixia	228	24	85.7	360	1	RUVB_COREG	Q8fpx5_corynebacte
156	24	85.7	221	2	Q554V9_DICTOST	Q554v9_dictyosteli	229	24	85.7	360	1	RUVB_COREG	Q8fpx5_corynebacte
157	24	85.7	221	2	Q554V9_DICTOST	Q554v9_dictyosteli	230	24	85.7	363	1	KCOL_ARATH	Q8l1b1_arabidopsais
158	24	85.7	221	2	Q4W494_AERSA	Q4w494_aeromonas s	231	24	85.7	369	2	Q49641_ARATH	Q49641_arabidopsais
159	24	85.7	221	2	Q599P2_AFRHY	Q599p2_aeromonas h	232	24	85.7	373	2	Q5Z200_NOCFA	Q5z200_nocardia fa
160	24	85.7	225	2	Q51L66_MAGGR	Q51l66_magnaporthe	233	24	85.7	374	2	Q6FUT0_CANGA	Q6fut0_candida gla
161	24	85.7	229	2	Q6AG89_LEIXX	Q6ag89_leifsonia x	234	24	85.7	375	2	Q9XG12_LYCES	Q9xg12_lycopersico
162	24	85.7	229	2	Q88A14_PSESM	Q88a14_pseudomonas	235	24	85.7	376	1	TRUC_ERWCA	Q74717_erwinia car
163	24	85.7	231	2	Q8NCV3_HUMAN	Q8ncv3_homo sapien	236	24	85.7	378	2	Q6YUS8_ORYSA	Q6yus8_oryza sativ
164	24	85.7	235	2	Q9PJH9_CHLMU	Q9pjh9_chlamydia m	237	24	85.7	381	2	Q5V294_HALMA	Q5v294_haloarcula
165	24	85.7	240	2	Q8E6F0_ANTILO	Q8e6f0_antonomospora	238	24	85.7	383	2	Q7U9U2_SYNXP	Q7u9u2_synechococc
166	24	85.7	245	2	Q662X5_BORGA	Q662x5_borrelia ga	239	24	85.7	384	2	Q88T33_LACPL	Q88t33_lactobacilli
167	24	85.7	246	2	Q7Q255_ANOGA	Q7q255_anopheles g	240	24	85.7	385	2	Q9LFA6_ARATH	Q9lfa6_arabidopsais
168	24	85.7	247	2	Q56BU1_9CAUD	Q56bu1_enterobacte	241	24	85.7	385	2	Q93YR5_ARATH	Q93yr5_arabidopsais
169	24	85.7	250	1	PLSC_BOREU	Q59188_borrelia bu	242	24	85.7	388	2	Q6NGG9_CORDI	Q6ngg9_corynebacte
170	24	85.7	252	1	SURE_RHILO	Q981c9_rhizobium l	243	24	85.7	403	1	Y4635_PSEPK	Q88e33_pseudomonas
171	24	85.7	254	2	Q8HN57_CRYNE	Q8hn57_cryptococcu	244	24	85.7	405	2	Q6CTO1_KIULA	Q6ct01_kluyveromyc
172	24	85.7	254	2	Q6J9U3_CRYNV	Q6j9u3_cryptococcu	245	24	85.7	407	2	Q8A6I2_PROAC	Q8a6i2_propanibac
173	24	85.7	254	2	Q8NMU6_CORGL	Q8nmu6_corynebacte	246	24	85.7	416	2	Q8ELK2_OCEIH	Q8elk2_oceanobacil
174	24	85.7	255	2	Q4K3B2_PSEFS	Q4k3b2_pseudomonas	247	24	85.7	423	2	Q8XQB6_CANAL	Q8xqb6_raistonia s
175	24	85.7	256	2	Q5JST5_HUMAN	Q5jst5_homo sapien	248	24	85.7	429	2	Q5A6B5_CANAL	Q5a6b5_candida alb
176	24	85.7	257	1	TRUC_YERPE	Q82h72_yersinia pe	249	24	85.7	431	2	Q6AM80_DESPS	Q6am80_desulfotale
177	24	85.7	257	2	Q667I3_YERPS	Q667i3_yersinia ps	250	24	85.7	432	2	Q5N2A4_SYNPE	Q5n2a4_synechococc

251	24	85.7	439	2	QBEDJ1 SHEON	Q8edj1 shewanella	324	24	85.7	650	2	Q4RF57_TETNG	Q4rf57 tetraodon n
252	24	85.7	446	2	Q61TK1_CAEBR	Q61tk1 caenorhabdi	325	24	85.7	651	2	Q7U3Y4_SYNXP	Q7u3y4 synchococc
253	24	85.7	448	2	Q7PM12_ANOGA	Q7pm12 anopheles g	326	24	85.7	654	1	HSP77_YEAST	P12398 saccharomyc
254	24	85.7	448	2	Q9GZH2_CAEEL	Q9gzh2 caenorhabdi	327	24	85.7	654	2	Q8Y0H3_RALSO	Q8y0h3 ralistonia s
255	24	85.7	452	2	Q5BGK3_ENENI	Q5bgk3 aspergillus	328	24	85.7	659	2	Q554W2_DICDI	Q554w2 dictyosteli
256	24	85.7	453	1	TOLB_GLUOX	Q5fqb9 gluonobact	329	24	85.7	671	2	Q97209_LEIMA	Q97209 leishmania
257	24	85.7	459	2	Q32384_SYN7P	Q32384 synchococc	330	24	85.7	688	2	Q5W7N9_BACTU	Q5w7n9 bacillus th
258	24	85.7	459	2	Q5N3A0_SYN7P	Q5n3a0 synchococc	331	24	85.7	693	1	EFGL_BORGA	Q60y4 borrelia ga
259	24	85.7	462	1	SAHH_RSODE	Q9zns5 roseobacter	332	24	85.7	699	2	Q6NZX0_BRARE	Q6nzo0 brachydanio
260	24	85.7	465	1	C85A1_ARATH	Q9fms5 arabidopsis	333	24	85.7	701	2	Q74DM6_GEOSL	Q74dm6 geobacter s
261	24	85.7	465	1	C85A2_ARATH	Q9fms5 arabidopsis	334	24	85.7	710	1	TRIM9_HUMAN	Q91zy8 rattus norv
262	24	85.7	466	2	Q7PNH9_ANOGA	Q9fms5 arabidopsis	335	24	85.7	710	1	Q4REG3_TETNG	Q4reg3 tetraodon n
263	24	85.7	467	2	Q6CT29_KJULA	Q6ct29 anopheles g	336	24	85.7	724	2	Q4RDX3_STTRI	Q4rdx3 streptomyc
264	24	85.7	469	2	Q06920_YEAST	Q6ct29 kluyveromyc	337	24	85.7	724	2	Q7SDQ6_NEUCR	Q7sdq6 neurospora
265	24	85.7	469	2	Q4WKH9_ASPFU	Q06920 saccharomyc	338	24	85.7	731	2	Q7UW15_RHOBA	Q7uwi5 rhodopirell
266	24	85.7	469	2	Q5DCR4_SCHJA	Q4wk9 aspergillus	339	24	85.7	733	2	Q4KF75_PSEF5	Q4kf75 pseudomonas
267	24	85.7	470	2	Q5JST2_HUMAN	Q5dcr4 schistosoma	340	24	85.7	745	2	Q8LJ22_SORBI	Q8lj22 sorghum bic
268	24	85.7	471	1	RUVB2_YEAST	Q5jst2 homo sapien	341	24	85.7	750	2	Q6G162_BARQU	Q6g162 bartonella
269	24	85.7	476	2	Q6FSF1_CANGA	Q12464 saccharomyc	342	24	85.7	754	2	Q6G326_BARHE	Q6g326 bartonella
270	24	85.7	476	2	Q8BKT9_MOUSE	Q8bkt9 mus musculu	343	24	85.7	754	2	Q6G326_BARHE	Q6g326 bartonella
271	24	85.7	480	2	Q8BSB8_DBBHA	Q8bkt9 mus musculu	344	24	85.7	762	2	Q7S9Z5_NEUCR	Q7s9z5 neurospora
272	24	85.7	481	2	Q5V3K3_DROME	Q8bkt9 mus musculu	345	24	85.7	781	2	Q8U9K3_AGR15	Q8u9k3 agrobacteri
273	24	85.7	482	2	Q8WSG2_9EUKA	Q8bkt9 mus musculu	346	24	85.7	788	2	Q7CT60_AGR15	Q7ct60 agrobacteri
274	24	85.7	482	2	Q5AGZ9_CANAL	Q8bkt9 mus musculu	347	24	85.7	791	2	Q73U76_MYCPA	Q73u76 mycobacteri
275	24	85.7	498	2	Q4TFW5_GIBZE	Q4fws5 gibberella	348	24	85.7	799	2	Q88F20_PSEPK	Q88f20 pseudomonas
276	24	85.7	501	2	BR22_YEAST	P41332 saccharomyc	349	24	85.7	806	2	Q6BRK5_DBBHA	Q6brk5 debaryomyc
277	24	85.7	505	1	BR22_YEAST	P41332 saccharomyc	350	24	85.7	813	2	Q9FJA1_ARATH	Q9fja1 arabidopsis
278	24	85.7	506	2	Q8EH17_SHEON	Q8eh17 shewanella	351	24	85.7	817	1	TRIM9_MOUSE	Q8c7m3 mus musculu
279	24	85.7	511	1	NIR_SYN7P	Q5BC23 aspergillus	352	24	85.7	831	2	Q5B3M2_EMENI	Q5b3m2 aspergillus
280	24	85.7	512	1	Q5N5B8_HUMAN	P39661 synchococc	353	24	85.7	842	2	Q82MV5_STRAW	Q82mv5 streptomyc
281	24	85.7	512	2	Q5TBD6_HUMAN	Q5n5b8 synchococc	354	24	85.7	843	2	Q4NMV7_9DELT	Q4nmv7 anaeromycob
282	24	85.7	513	2	Q5FIH2_ARATH	Q5fih2 homo sapien	355	24	85.7	843	2	Q72RNS_LEPIC	Q72rn5 leptospira
283	24	85.7	517	2	Q4UD07_THEAN	Q4ud07 theileria a	356	24	85.7	884	2	Q5GU97_XANOR	Q5gu97 xanthomonas
284	24	85.7	521	2	Q98C54_RHILO	Q98c54 rhizobium l	357	24	85.7	924	2	Q59Q22_CANAL	Q59q22 candida alb
285	24	85.7	527	2	Q9DW52_RCMVM	Q9dw52 rat cytomeg	358	24	85.7	936	2	Q4RJ27_TETNG	Q4rj27 tetraodon n
286	24	85.7	530	2	Q7O4U7_ANOGA	Q7o4u7 anopheles g	359	24	85.7	952	2	Q7XN31_ORYSA	Q7xn31 oryza sativ
287	24	85.7	535	2	Q95K16_MACFA	Q95k16 macaca fasc	360	24	85.7	960	2	Q7N7N1_PHOLL	Q7n7n1 photorhabdu
288	24	85.7	538	2	Q983Q8_RHILO	Q983q8 rhizobium l	361	24	85.7	966	2	Q4NMD6_9MICC	Q4nmd6 arthrobacte
289	24	85.7	541	2	Q983Q8_RHILO	Q983q8 rhizobium l	362	24	85.7	988	2	Q600R9_MYCHY	Q600r9 mycoplasma
290	24	85.7	543	2	Q8E5B1_STRA3	Q8e5b1 streptococc	363	24	85.7	993	2	Q9U4F8_DROME	Q9u4f8 drosophila
291	24	85.7	551	1	KIF12_HUMAN	Q8e5b1 streptococc	364	24	85.7	997	2	Q9GYL4_CAEEL	Q9gy14 caenorhabdi
292	24	85.7	551	2	Q7UWH7_RHOBA	Q7uwh7 rhodopirell	365	24	85.7	1000	1	Y041_SYN73	Q55445 synchocyst
293	24	85.7	564	2	Q7S0Z2_NEUCR	Q7s0z2 neurospora	366	24	85.7	1000	2	Q4A855_SYN73	Q55445 synchocyst
294	24	85.7	564	2	Q4KCV3_PSEF5	Q4kcv3 pseudomonas	367	24	85.7	1027	2	Q695H0_CHLRE	Q695h0 chlamydomon
295	24	85.7	568	2	Q4L5N0_PSEAE	Q4l5n0 pseudomonas	368	24	85.7	1061	2	Q4XQ34_PLACH	Q4xq34 plasmodium
296	24	85.7	569	2	Q98KA7_RHILO	Q98ka7 rhizobium l	369	24	85.7	1061	2	Q8RNY3_HAPACH	Q8rny3 hafnia alve
297	24	85.7	573	2	Q4SFK8_TETNG	Q4sfk8 tetraodon n	370	24	85.7	1070	2	Q615X9_ORYSA	Q615x9 oryza sativ
298	24	85.7	585	2	Q9V3M9_DROME	Q9v3m9 drosophila	371	24	85.7	1087	2	Q4HWP8_GIBZE	Q4hwp8 gibberella
299	24	85.7	599	2	Q7NYB6_CHRVO	Q7nyb6 chromobacte	372	24	85.7	1103	2	Q7V6Q7_PROMM	Q7v6q7 prochloroco
300	24	85.7	600	2	Q9HSM3_HALSA	Q9hsm3 halobacteri	373	24	85.7	1104	2	Q9XEQ2_SORBI	Q9xeq2 sorghum bic
301	24	85.7	600	2	Q4PE22_USTWA	Q4pe22 ustilago ma	374	24	85.7	1105	2	Q9HWT9_GIBZE	Q9hwt9 gibberella
302	24	85.7	600	2	Q869TE_DICDI	Q869te dictyosteli	375	24	85.7	1140	2	Q4WNA6_ASPFU	Q4wna6 aspergillus
303	24	85.7	600	2	Q5HTL6_CAMJR	Q5htl6 campylobact	376	24	85.7	1175	2	Q92PF3_ARATH	Q92pf3 arabidopsis
304	24	85.7	600	2	Q9PN51_CAMJR	Q9pn51 campylobact	377	24	85.7	1175	2	Q9SMU3_ARATH	Q9smu3 arabidopsis
305	24	85.7	600	2	Q8E4V0_MYCPE	Q8e4v0 mycoplasma	378	24	85.7	1179	2	Q7SAB3_NEUCR	Q7sab3 neurospora
306	24	85.7	602	2	Q8F426_LEPIN	Q8f426 leptospira	379	24	85.7	1195	2	Q4ICQ6_GIBZE	Q4icq6 gibberella
307	24	85.7	603	2	Q9L217_STRCO	Q9l217 streptomyc	380	24	85.7	1203	2	Q9SN55_ARATH	Q9sn55 arabidopsis
308	24	85.7	604	2	Q55F14_DICDI	Q55f14 dictyosteli	381	24	85.7	1269	2	Q7UH58_RHOBA	Q7uh58 rhodopirell
309	24	85.7	606	2	Q4TR89_9SPHN	Q4tr89 erythrobact	382	24	85.7	1326	1	CI093_HUMAN	Q6tfl3 homo sapien
310	24	85.7	610	2	Q6GLX6_XENLA	Q6glx6 xenopus lae	383	24	85.7	1358	2	Q6L3H0_SOLDE	Q6l3h0 solanum dem
311	24	85.7	615	2	Q512R9_ENTHI	Q512r9 entamoeba h	384	24	85.7	1484	2	Q5A6J4_CANAL	Q5a6j4 candida alb
312	24	85.7	621	2	Q7YXJ6_TRYCR	Q7yxj6 trypanosoma	385	24	85.7	1489	2	Q23302_ARATH	Q23302 arabidopsis
313	24	85.7	624	2	Q8GSP3_BIFLO	Q8gsp3 bifidobacte	386	24	85.7	1545	2	Q5KKG1_CRYNE	Q5kkg1 cryptococcu
314	24	85.7	628	1	FTSH3_SYN73	Q753p3 synchocyst	387	24	85.7	1545	2	Q81617_ARATH	Q81617 arabidopsis
315	24	85.7	628	2	Q66DT0_YERPS	Q66dt0 yersinia ps	388	24	85.7	1633	2	Q7PS39_ANOGA	Q7ps39 anopheles g
316	24	85.7	628	2	Q8ZC69_YERPE	Q8zcc69 yersinia pe	389	24	85.7	1702	2	Q59DT7_DROME	Q59dt7 drosophila
317	24	85.7	634	2	Q95JX7_MACFA	Q95jx7 macaca fasc	390	24	85.7	1756	2	Q6AWK8_DROME	Q6awk8 drosophila
318	24	85.7	640	2	Q5TKI6_ORYSA	Q5tki6 oryza sativ	391	24	85.7	1850	2	Q59DT8_DROME	Q59dt8 drosophila
319	24	85.7	646	2	Q5TBE0_HUMAN	Q5tbe0 homo sapien	392	24	85.7	1851	1	LAPA_DROME	Q7kyl5 plasmodium
320	24	85.7	646	2	Q8JH68_BRARE	Q8jh68 brachydanio	393	24	85.7	1851	2	Q7RFL5_PLAYO	Q7rf15 cherry gree
321	24	85.7	648	2	Q9HUH9_PSEAE	Q9huh9 pseudomonas	394	24	85.7	1986	2	Q91632_9VIRU	Q91632 cherry gree
322	24	85.7	650	2	Q4KD73_PSEF5	Q4kd73 pseudomonas	395	24	85.7	2023	2	Q7N7K0_PHOLL	Q7n7k0 photorhabdu
323	24	85.7	650	2	Q88IY0_PSEPK	Q88iy0 pseudomonas	396	24	85.7	2937	2		

397	24	85.7	2961	2	Q7MB60	PHOLL	Q7mb60	photorhabdu	470	23	82.1	204	2	Q4T3U3	TETNG	Q4t3u3	tetradon n
398	24	85.7	3027	2	Q7MB03	PHOLL	Q7mb03	photorhabdu	471	23	82.1	205	2	Q5V7J6	HALMA	Q5v7j6	haloarculla
399	24	85.7	3056	1	POLG	BYMV	P17765	b. genome po	472	23	82.1	205	2	Q889C8	PSESM	Q889c8	pseudomonas
400	24	85.7	3056	2	Q58852	BYMV	Q58852	bean yellow	473	23	82.1	209	2	Q5NFZ1	FRATT	Q5nfz1	francisella
401	24	85.7	3056	2	Q58892	BYMV	Q58892	bean yellow	474	23	82.1	211	2	Q8X147	ALTR	Q8x147	alternaria
402	24	85.7	3072	2	Q92645	CVVV	Q92645	clover yell	475	23	82.1	211	2	Q98EE8	RHILO	Q98ee8	rhizobium 1
403	24	85.7	3501	2	Q8Y106	RALSO	Q8y106	raistonia s	476	23	82.1	212	2	Q5WZS5	LEGPL	Q5wzs5	legionella
404	24	85.7	3552	2	Q8XSD6	RALSO	Q8xsd6	raistonia s	477	23	82.1	212	2	Q5X8D0	LEGPA	Q5x8d0	legionella
405	24	85.7	4582	2	Q7N911	PHOLL	Q7n911	photorhabdu	478	23	82.1	212	2	Q8FLT7	COREF	Q8flt7	corynebacte
406	24	85.7	4602	2	Q6V7I9	9CAUD	Q6v7i9	burkholderi	479	23	82.1	213	2	Q8DM96	SYNEL	Q8dm96	synecococc
407	24	85.7	7538	2	Q7UNV3	RHOBA	Q7unv3	rhodopirell	480	23	82.1	216	1	MTTC	METBA	Q93659	methanogarc
408	23	82.1	60	2	Q9TX85	CAEEL	Q9tx85	caenotrhabi	481	23	82.1	216	2	Q6FFI8	ACIAD	Q6ffi8	acinetoabact
409	23	82.1	61	2	Q8J152	9PLEO	Q8j152	leptosphaer	482	23	82.1	218	1	MTTC2	METMA	P58983	methanosarc
410	23	82.1	83	1	GVPG2	HALSA	P33960	halobacteri	483	23	82.1	220	2	Q5WTD8	LEGPL	Q5wtd8	legionella
411	23	82.1	89	2	Q4RCZ2	TETNG	Q4rcz2	tetradon n	484	23	82.1	220	2	Q5WVJ8	LEGPL	Q5wvj8	legionella
412	23	82.1	91	1	H530E	XENLA	P30220	xenopus lae	485	23	82.1	220	2	Q5XIM9	LEGPA	Q5xim9	legionella
413	23	82.1	93	2	Q88BF2	PSEPK	Q88bf2	pseudomonas	486	23	82.1	220	2	Q5X462	LEGPA	Q5x462	legionella
414	23	82.1	95	2	Q4FLO6	GRICK	Q4flq6	caudatus	487	23	82.1	220	2	Q5Z859	LEGPH	Q5z859	legionella
415	23	82.1	97	2	Q65LJ0	BACLD	Q65lj0	bacillus li	488	23	82.1	220	2	Q5ZUE7	LEGPH	Q5zue7	legionella
416	23	82.1	102	2	Q71I95	LACDL	Q71ifs	lactobacill	489	23	82.1	221	2	Q919T2	9INFA	Q919t2	influenza a
417	23	82.1	102	2	Q60B40	METCA	Q60b40	methylcoccc	490	23	82.1	221	2	Q919T3	9INFA	Q919t3	influenza a
418	23	82.1	110	2	Q60C52	METCA	Q60c52	methylcoccc	491	23	82.1	222	2	Q7NCM0	GLOVI	Q7ncm0	gloeobacter
419	23	82.1	115	2	Q8E187	SHEON	Q8e187	shewanella	492	23	82.1	226	2	Q5YWM5	NOCFA	Q5ywm5	nocardia fa
420	23	82.1	119	2	Q5UZV1	HALMA	Q5uzv1	haloarculla	493	23	82.1	228	2	Q81PL3	BACAN	Q81pl3	bacillus an
421	23	82.1	121	2	Q6FCV2	ACIAD	Q6fcv2	acinetoabact	494	23	82.1	228	2	Q6HHU8	BACHK	Q6hhu8	bacillus th
422	23	82.1	122	1	RL18	RHOBA	Q7un04	rhodopirell	495	23	82.1	228	2	Q7UH08	RHOBA	Q7uh08	rhodopirell
423	23	82.1	129	2	Q825V1	STRAW	Q825v1	streptomyce	496	23	82.1	233	2	Q6UE31	9LACO	Q6ue31	lactobacill
424	23	82.1	132	2	Q9AN59	BRAJA	Q9an59	bradyrhizob	497	23	82.1	236	2	Q74886	SCHPO	Q74886	schistosacch
425	23	82.1	135	2	Q8TPW1	METAC	Q8tpw1	methanosearc	498	23	82.1	237	2	Q4QAP9	LEIMA	Q4qap9	leishmania
426	23	82.1	140	2	Q9SAL1	ARATH	Q9sal1	arabidopsais	499	23	82.1	238	2	Q81922	CAPAN	Q81922	capsicum an
427	23	82.1	143	2	Q87FW9	VIBPA	Q87fw9	vibrio para	500	23	82.1	241	2	Q8RM80	LEGPN	Q8rm80	legionella
428	23	82.1	143	2	Q6SUZ8	9BACU	Q6su28	achaea jana	501	23	82.1	242	2	Q604M9	METCA	Q604m9	methylococc
429	23	82.1	144	2	Q8K3C6	ORYSA	Q8k3c6	oryza sativ	502	23	82.1	242	2	Q87EB8	XYLFT	Q87ee8	xylella fas
430	23	82.1	147	2	Q825J2	STRAW	Q825j2	streptomyce	503	23	82.1	245	2	Q4NUG0	9DELT	Q4nug0	anaeromyxob
431	23	82.1	148	2	Q8PJ88	XANAC	Q8pj88	xanthomonas	504	23	82.1	248	2	Q05077	STRCL	Q05077	streptomyce
432	23	82.1	149	2	Q8ZH71	YERPE	Q8zh71	yersinia ps	505	23	82.1	249	1	UBIG	RHILO	Q98987	rhizobium 1
433	23	82.1	149	2	Q66714	YERPE	Q66714	yersinia ps	506	23	82.1	251	2	Q32722	PAEPO	Q32722	paenibacill
434	23	82.1	150	2	Q508Z8	ENTHI	Q508z8	entamoeba h	507	23	82.1	262	2	Q5Z8E2	ORYSA	Q5z8e2	oryza sativ
435	23	82.1	150	2	Q66086	9ALPH	Q66086	canine herp	508	23	82.1	263	2	Q8PPV1	XANAC	Q8ppv1	xanthomonas
436	23	82.1	151	2	Q87GT1	VIBPA	Q87gt1	vibrio para	509	23	82.1	264	2	Q66V92	9ACTO	Q66v92	nonomurea
437	23	82.1	153	2	Q4H6G1	9DEIO	Q4h6g1	deinococcus	510	23	82.1	265	2	Q5WMK2	BACSK	Q5wmk2	bacillus cl
438	23	82.1	154	1	PFDA	AERPE	Q9yd28	aeropyrum p	511	23	82.1	265	2	Q7U2P1	MYCBO	Q7u2p1	mycobacteri
439	23	82.1	158	2	Q5YS59	NOCFA	Q5ys89	nocardia fa	512	23	82.1	265	2	Q07412	MYCTO	Q07412	mycobacteri
440	23	82.1	161	2	Q6I896	XENLA	Q6i896	xenopus lae	513	23	82.1	268	2	Q9AEZ0	9ACTO	Q9aez0	frankia sp.
441	23	82.1	167	2	Q8V1N8	9PSED	Q8v1n8	pseudomonas	514	23	82.1	269	2	Q5V0B1	HALMA	Q5v0b1	haloarculla
442	23	82.1	167	2	Q6G502	BARHE	Q6g502	bartonella	515	23	82.1	269	2	Q4SKD7	TETNG	Q4skd7	tetradon n
443	23	82.1	167	2	Q8ELM3	OCEIH	Q8elm3	oceanobacil	516	23	82.1	275	2	Q9HDD4	9PLEO	Q9hdd4	alternaria
444	23	82.1	168	2	Q4LZK2	9BURK	Q4lzk2	burkholderi	517	23	82.1	275	2	Q9HES6	9PLEO	Q9hes6	alternaria
445	23	82.1	169	2	Q7Q9Y1	ANOGA	Q7q9y1	anopheles g	518	23	82.1	275	2	Q7UZC5	RHOBA	Q7uzc5	rhodopirell
446	23	82.1	169	2	Q50NW5	ENTHI	Q50nw5	entamoeba h	519	23	82.1	278	2	Q8X146	9PLEO	Q8x146	alternaria
447	23	82.1	170	2	Q5UQC6	MIMIV	Q5uqc6	chlamydia p	520	23	82.1	278	2	Q8BCC5	9INFA	Q8bcc5	influenza a
448	23	82.1	171	2	Q9Z7K5	CHLPN	Q9z7k5	chlamydia p	521	23	82.1	278	2	Q8BCC6	9INFA	Q8bcc6	influenza a
449	23	82.1	172	2	Q5L767	CHLAB	Q5l767	chlamydophi	522	23	82.1	278	2	Q8BCC7	9INFA	Q8bcc7	influenza a
450	23	82.1	172	2	Q824U7	CHLCV	Q824u7	chlamydophi	523	23	82.1	278	2	Q8BCC8	9INFA	Q8bcc8	influenza a
451	23	82.1	173	2	Q846B3	CHLTR	Q846b3	chlamydia t	524	23	82.1	278	2	Q8BCC9	9INFA	Q8bcc9	influenza a
452	23	82.1	174	2	Q7V672	PROMM	Q7v672	prochloroco	525	23	82.1	278	2	Q8BCD0	9INFA	Q8bcd0	influenza a
453	23	82.1	177	2	Q9PLP8	CHLMU	Q9plp8	chlamydia m	526	23	82.1	278	2	Q8BCD1	9INFA	Q8bcd1	influenza a
454	23	82.1	180	2	Q9PAL1	XYLFA	Q9pal1	xylella fas	527	23	82.1	278	2	Q8BCD2	9INFA	Q8bcd2	influenza a
455	23	82.1	180	2	Q67JN3	SYMTH	Q67jn3	syntrophobacte	528	23	82.1	278	2	Q8BCD3	9INFA	Q8bcd3	influenza a
456	23	82.1	180	2	Q877M8	XYLFT	Q877m8	xylella fas	529	23	82.1	278	2	Q8BCD4	9INFA	Q8bcd4	influenza a
457	23	82.1	181	2	Q8DH08	SYNEL	Q8dh88	synecococc	530	23	82.1	283	2	Q5JW01	HUMAN	Q5jw01	homo sapien
458	23	82.1	184	2	Q65P08	BACLD	Q65pd8	bacillus li	531	23	82.1	289	2	Q5GMG5	MOUSE	Q5gmg5	mus muscucu
459	23	82.1	190	2	Q72QY6	LEPIC	Q72qy6	leptospira	532	23	82.1	291	2	Q7PSX0	ANOGA	Q7psx0	anopheles g
460	23	82.1	190	2	Q8F4V4	LEPIN	Q8f4v4	leptospira	533	23	82.1	297	2	Q86EA5	SCHJA	Q86ea5	schistosoma
461	23	82.1	192	1	Y4GC	RHSIN	P55459	rhizobium s	534	23	82.1	297	2	Q88LC6	PSEPK	Q88lc6	pseudomonas
462	23	82.1	193	2	Q859F3	9VIRU	Q859f3	uncultured	535	23	82.1	298	2	Q6U737	VIBFI	Q6u737	vibrio fiac
463	23	82.1	193	2	Q8EDR6	SHEON	Q8edr6	shewanella	536	23	82.1	301	2	Q24760	HYPME	Q24760	hyphomicrob
464	23	82.1	198	2	Q859N6	9VIRU	Q859n6	uncultured	537	23	82.1	302	2	Q6LS86	PHOPR	Q6ls86	photobacter
465	23	82.1	203	2	Q32C91	LASIN	Q32c91	listeria in	538	23	82.1	302	2	Q88YC3	LACPL	Q88yc3	lactobacill
466	23	82.1	204	2	Q7PE10	ANOGA	Q7pe10	anopheles g	539	23	82.1	303	1	FLSC	FEAST	P33333	saccharomyc
467	23	82.1	204	2	Q4ZUD7	PSESY	Q4zud7	pseudomonas	540	23	82.1	303	2	Q6B238	YEAST	Q6b238	saccharomyc
468	23	82.1	204	2	Q8EX26	MYCPE	Q8ex26	mycoplasma	541	23	82.1	303	2	Q97147	CLOAB	Q97147	clostridium
469	23	82.1	204	2	Q6D313	ERWCT	Q6d313	erwinia car	542	23	82.1	306	2	Q9W5G7	DROME	Q9w5g7	drosophila

543	23	82.1	308	2	Q9HRV7 HALSA	Q9hrv7 halobacteri	616	23	82.1	379	2	Q69FI2_9PLEO	Q69fi2 stemphylium
544	23	82.1	308	2	Q8DFI3_VIBVU	Q8dfi3 vibrio vuln	617	23	82.1	379	2	Q69FI4_9PLEO	Q69fi4 stemphylium
545	23	82.1	308	2	Q7MMV9_VIBVU	Q7mmv9 vibrio vuln	618	23	82.1	379	2	Q69FI6_9PLEO	Q69fi6 stemphylium
546	23	82.1	309	1	METE_METTM	P55299 methanobact	619	23	82.1	379	2	Q69FI8_9PLEO	Q69fi8 stemphylium
547	23	82.1	312	2	Q5NBE6_ORISA	Q5nbe6 oryza sativ	620	23	82.1	379	2	Q69FJ0_9PLEO	Q69fj0 stemphylium
548	23	82.1	313	2	Q8K3P0_MOUSE	Q8k3p0 m retinol d	621	23	82.1	379	2	Q69FJ2_9PLEO	Q69fj2 stemphylium
549	23	82.1	315	2	Q7NLP1_GLOVI	Q7nlp1 gloeobacter	622	23	82.1	379	2	Q69FJ4_9PLEO	Q69fj4 stemphylium
550	23	82.1	315	2	Q7UHR2_RHOBA	Q7uhr2 rhodopirell	623	23	82.1	379	2	Q69FJ6_9PLEO	Q69fj6 pleospora s
551	23	82.1	315	2	Q6WVD8_MOUSE	Q6wvd8 mus musculu	624	23	82.1	379	2	Q69FJ9_9PLEO	Q69fj9 pleospora s
552	23	82.1	316	2	Q6N9F4_RHOPA	Q6n9f4 rhodopseudo	625	23	82.1	379	2	Q69G97_9PLEO	Q69g97 pleospora s
553	23	82.1	317	1	YVCK_BACSU	O06974 bacillus su	626	23	82.1	379	2	Q69GA0_9PLEO	Q69ga0 stemphylium
554	23	82.1	317	2	Q518K9_ENTHI	Q518k9 entamoeba h	627	23	82.1	379	2	Q69GA3_9PLEO	Q69ga3 pleospora s
555	23	82.1	317	2	Q5FHM6_EHRRG	Q5fhm6 ehrlichia r	628	23	82.1	379	2	Q69GA6_9PLEO	Q69ga6 pleospora p
556	23	82.1	317	2	Q5YCR2_BRUME	Q5ycr2 bruceella me	629	23	82.1	379	2	Q69GA9_9PLEO	Q69ga9 pleospora e
557	23	82.1	317	2	Q5HBJ7_EHRRW	Q5hbj7 ehrlichia r	630	23	82.1	379	2	Q69GB2_9PLEO	Q69gb2 pleospora t
558	23	82.1	319	2	Q5WK78_LEGPL	Q5wk78 legionella	631	23	82.1	379	2	Q69GB5_9PLEO	Q69gb5 stemphylium
559	23	82.1	319	2	Q5XSV6_LEGPA	Q5xsv6 legionella	632	23	82.1	379	2	Q69GC1_9PLEO	Q69gc1 pleospora s
560	23	82.1	321	2	Q9SSA3_DROME	Q9ssea3 drosophila	633	23	82.1	379	2	Q69GC4_9PLEO	Q69gc4 stemphylium
561	23	82.1	321	2	Q9U602_ANISI	Q9u602 anisakis si	634	23	82.1	379	2	Q69GC7_9PLEO	Q69gc7 stemphylium
562	23	82.1	321	2	Q9V8J9_DROME	Q9v8j9 drosophila	635	23	82.1	379	2	Q69GD0_9PLEO	Q69gd0 stemphylium
563	23	82.1	323	2	Q5ZW75_LEGPH	Q5zw75 legionella	636	23	82.1	379	2	Q69GD3_9PLEO	Q69gd3 pleospora g
564	23	82.1	324	2	Q4ISP7_AZOVI	Q4isp7 azotobacter	637	23	82.1	379	2	Q69GD6_9PLEO	Q69gd6 pleospora t
565	23	82.1	328	2	Q9HSV1_HALSA	Q9hsv1 halobacteri	638	23	82.1	379	2	Q69GD9_9PLEO	Q69gd9 pleospora a
566	23	82.1	328	2	Q5YMS7_NOCFA	Q5yms7 nocardia fa	639	23	82.1	379	2	Q69GE2_9PLEO	Q69ge2 pleospora h
567	23	82.1	329	2	Q6BTZ6_DEBHA	Q6btz6 debaryomyce	640	23	82.1	379	2	Q69GE5_9PLEO	Q69ge5 pleospora t
568	23	82.1	329	2	Q7WEA3_BORBR	Q7wea3 bordetella	641	23	82.1	380	2	Q8VVI0_ERMPY	Q8vvi0 erwinia pyr
569	23	82.1	331	2	Q6GPM5_XENLA	Q6gpm5 xenopus lae	642	23	82.1	380	2	Q9R6M5_9RHIZ	Q9r6m5 agrobacteri
570	23	82.1	331	2	Q6P340_XENTR	Q6p340 xenopus tro	643	23	82.1	381	2	Q5LS50_SILPO	Q5ls50 silicibacte
571	23	82.1	332	1	LDHD_LACDE	Q8dq19 vibrio vuln	644	23	82.1	381	2	Q59311_CLOPE	Q59311 clostridium
572	23	82.1	332	1	Q8DG19_VIBVU	Q8dg19 vibrio vuln	645	23	82.1	382	2	Q9Z4A5_9ZZZZ	Q9z4a5 plasamid col
573	23	82.1	334	2	Q5UYL1_HALMA	Q5uy11 haloarcula	647	23	82.1	382	2	Q79VT3_SALTY	Q79vt3 salmonella
574	23	82.1	334	2	Q9KGG5_BACHD	Q9kgs5 bacillus ha	648	23	82.1	382	2	Q4FIH2_9ENTR	Q4fih2 salmonella
575	23	82.1	336	1	LDHD_LACHE	P30901 lactobacill	649	23	82.1	383	1	Q7DJM3_9ZZZZ	Q7djm3 plasmid r64
576	23	82.1	336	1	Q9ZH60_LACJO	Q9zh60 lactobacill	650	23	82.1	383	1	MAT1_COEHE	Q02990 cochllobolu
577	23	82.1	337	2	Q67K15_SYMTH	Q67k15 symbiobacte	651	23	82.1	384	2	Q4SUA9_TETNG	Q4sua9 tetraodon n
578	23	82.1	338	2	Q8IKK4_PLAF7	Q8ikk4 plasmodium	652	23	82.1	384	2	Q5WKE3_BACSK	Q5wke3 bacillus cl
579	23	82.1	344	2	Q9GR09_PLAFA	Q9gr09 plasmodium	653	23	82.1	385	2	Q5SQ48_CRYNE	Q5sq48 cryptococcu
580	23	82.1	344	2	Q5FNM5_LACAC	Q5fnm5 lactobacill	654	23	82.1	385	2	Q5KDC0_CRYNE	Q5kdc0 cryptococcu
581	23	82.1	349	2	Q6F7F1_ACIAD	Q6f7f1 acinetobact	655	23	82.1	387	1	S14L2_BOVIN	P58875 bos taurus
582	23	82.1	349	2	Q6VYQ3_ANASP	Q6vynq3 anabaena ep	656	23	82.1	388	2	Q4HBY9_9DEIO	Q4hby9 deinococcus
583	23	82.1	351	1	MURB_TROW8	Q83ha7 tropheryma	657	23	82.1	388	2	Q8XK16_BRAJA	Q8xk16 bradyrhizob
584	23	82.1	351	1	MURB_TROW8	Q83fk5 tropheryma	658	23	82.1	388	2	Q8XK16_BRAJA	Q8xk16 bradyrhizob
585	23	82.1	351	1	MURB_TROW8	Q83fk5 tropheryma	659	23	82.1	389	1	MAT1_ALTAL	Q94160 alternaria
586	23	82.1	354	2	Q7V896_PROMM	Q7v896 prochloroco	660	23	82.1	390	2	Q8FBQ7_ECOL6	Q8fbq7 escherichia
587	23	82.1	355	2	Q5FR15_GLUOX	Q5fr15 gluconobact	661	23	82.1	390	2	Q83PI2_SHIFL	Q83pi2 shigella fl
588	23	82.1	355	2	Q8G4J8_BIFLO	Q8g4j8 bifidobacte	662	23	82.1	391	2	Q84ES3_9BURK	Q84es3 cupriavidus
589	23	82.1	356	2	Q501X3_BRARE	Q501x3 brachydanio	663	23	82.1	392	2	Q6PD61_HUMAN	Q6pd61 homo sapien
590	23	82.1	357	1	CV008_MOUSE	Q91vnl mus musculu	664	23	82.1	392	2	Q5REK6_PONPY	Q5rek6 pongo pygma
591	23	82.1	360	2	Q87CI2_XYLFT	Q87ci2 xylella fas	665	23	82.1	392	2	Q9RW09_DEIRA	Q9rw09 deinococcus
592	23	82.1	360	2	Q9PAL7_XYLPA	Q9pal7 xylella fas	666	23	82.1	392	2	Q9CW5_MOUSE	Q9cwm5 mus musculu
593	23	82.1	362	1	RUVB_CORDI	Q5czms brachydanio	667	23	82.1	394	2	Q59GA6_HUMAN	Q59ga6 homo sapien
594	23	82.1	362	1	RUVB_CORDI	P61530 corynebacte	668	23	82.1	398	2	Q4H345_CIOIN	Q4h345 ciona intes
595	23	82.1	363	2	Q72V50_LEPIC	Q72v50 leptospira	669	23	82.1	399	2	Q88TC7_LACPL	Q88tc7 lactobacill
596	23	82.1	363	2	Q8EQ08_LEPIN	Q8eq08 leptospira	670	23	82.1	400	1	S14L3_HUMAN	Q9udx4 homo sapien
597	23	82.1	364	2	Q6FXX6_CANGA	Q6fxx6 candida gla	671	23	82.1	400	1	S14L3_RAT	Q9z1j8 rattus norv
598	23	82.1	364	2	Q5FN37_GLUOX	Q5fn37 gluconobact	672	23	82.1	400	1	S14L3_RAT	Q9z1j8 rattus norv
599	23	82.1	365	2	Q987N7_RHILO	Q987n7 rhizobium l	673	23	82.1	400	2	Q6XCI7_HUMAN	Q6xc17 homo sapien
600	23	82.1	367	2	Q91V19_MOUSE	Q91v19 mus musculu	674	23	82.1	401	2	Q61SB2_HUMAN	Q61sb2 homo sapien
601	23	82.1	368	2	Q740G9_MYCPA	Q740g9 mycobacteri	675	23	82.1	403	1	Q5SQ27_MOUSE	Q5sq27 mus musculu
602	23	82.1	369	2	Q5Y867_NOCFA	Q5ys87 nocardia fa	676	23	82.1	403	1	S14L2_MOUSE	Q76054 homo sapien
603	23	82.1	369	2	Q6AER0_LEIXX	Q6aer0 leifsonia x	677	23	82.1	403	2	S14L2_MOUSE	Q99J08 mus musculu
604	23	82.1	373	2	Q6BF17_PARTE	Q6bf17 paramecium	678	23	82.1	403	2	Q53EQ2_HUMAN	Q99m80 rattus norv
605	23	82.1	376	1	WECB_ECOLI	Q8xar8 escherichia	679	23	82.1	403	2	Q67A0_BOVIN	Q5sq25 mus musculu
606	23	82.1	376	1	WECB_ECOLI	P27828 escherichia	680	23	82.1	403	2	Q5EBD0_RAT	Q5ebd0 rattus norv
607	23	82.1	377	2	Q5L2A6_GEOKA	Q5l2a6 geobacillus	681	23	82.1	404	2	Q75W78_BACAM	Q75w78 brassica ca
608	23	82.1	378	1	MAT1_COCCA	Q9p446 cochllobolu	682	23	82.1	404	2	Q82W26_LEPIC	Q82w26 leptospira
609	23	82.1	379	1	MAT1_COCCA	O13402 cochllobolu	683	23	82.1	404	2	Q8F9R4_LEPIN	Q8f9r4 leptospira
610	23	82.1	379	1	MAT1_COCCA	Q9y8d3 cochllobolu	684	23	82.1	405	2	Q9KQ11_VIBCH	Q9kg11 vibrio chol
611	23	82.1	379	1	MAT1_COCCA	Q9y8d3 cochllobolu	685	23	82.1	408	2	Q7R536_SHIFL	Q7rs36 giardia lam
612	23	82.1	379	2	Q69FH2_9PLEO	Q69fh2 stemphylium	686	23	82.1	408	2	Q83PX2_SHIFL	Q83px2 shigella fl
613	23	82.1	379	2	Q69FH4_9PLEO	Q69fh4 stemphylium	687	23	82.1	409	1	AMAB1_BACST	P73113 bacillus st
614	23	82.1	379	2	Q69FH6_9PLEO	Q69fh6 pleospora s	688	23	82.1	409	1	AMAB2_BACST	Q53389 bacillus st
615	23	82.1	379	2	Q69FH8_9PLEO	Q69fh8 stemphylium							
					Q69f10 stemphylium								

689	23	82.1	409	2	O80791	arabidopsis	762	23	82.1	457	2	Q4NT81	9DELT	Q4nt81 anaeromyxob
690	23	82.1	409	2	O6YZ60	oryza sativ	763	23	82.1	457	2	O6DDN1	XENLA	Q6ddn1 xenopus lae
691	23	82.1	409	2	O8GQ95	geobacillus	764	23	82.1	459	2	Q74L28	LACJO	Q74l28 lactobacill
692	23	82.1	411	2	Q8QVC7	geoch virus	765	23	82.1	459	2	Q58EL7	BRARE	Q58el7 brachydanio
693	23	82.1	411	2	Q5RH89	brachydanio	766	23	82.1	460	1	RAB31	RAT	Q82739 rattus norv
694	23	82.1	413	2	Q9VJ17	drosofila	767	23	82.1	460	1	SYC	PSESM	Q87192 pseudomonas
695	23	82.1	413	2	Q9NK60	drosofila	768	23	82.1	460	2	O4K9S1	PSEF5	Q4k9s1 pseudomonas
696	23	82.1	415	2	Q4J4F0	azotovi	769	23	82.1	460	2	O4K9S1	PSEF5	Q4k9s1 pseudomonas
697	23	82.1	415	2	Q9J1B4	MOUSE	770	23	82.1	460	2	O8E9P4	SHEON	Q8e9p4 shewanella
698	23	82.1	418	2	Q72TP7	LEPIC	771	23	82.1	461	2	O8E9P4	SHEON	Q8e9p4 shewanella
699	23	82.1	418	2	Q8FLJ2	LEPIN	772	23	82.1	461	2	O8E9P4	SHEON	Q8e9p4 shewanella
700	23	82.1	420	2	Q4JX86	CORJCK	773	23	82.1	462	2	Q21794	CABEL	Q21794 caenorhabdi
701	23	82.1	423	2	Q75TW9	BACHD	774	23	82.1	465	2	O524T0	MAGGR	Q524t0 magnaporthe
702	23	82.1	424	2	Q4ZSN7	PSESY	775	23	82.1	465	2	O4MZ99	THEIPA	Q4mz99 theileria p
703	23	82.1	424	2	Q5L579	CHLAB	776	23	82.1	465	2	O7V3A5	PROMP	Q7v3a5 prochloroco
704	23	82.1	425	2	Q7WX91	ALCEU	777	23	82.1	465	2	O6P8B3	XENTR	Q6p8b3 xenopus tro
705	23	82.1	425	2	Q4M0Q4	9BURK	778	23	82.1	469	2	Q9L3C1	ANYMD	Q9l3c1 amycolatops
706	23	82.1	425	2	Q5QV89	IDILO	779	23	82.1	470	1	LEU2	BACSK	Q5wens bacillus cl
707	23	82.1	425	2	Q5O3R3	BRARE	780	23	82.1	470	2	O9GZ16	CABEL	Q9gz16 caenorhabdi
708	23	82.1	426	2	Q4UK60	RICKETTSIA	781	23	82.1	470	2	O7UIF9	MYCBO	Q7uif9 mycobacteri
709	23	82.1	426	2	Q68VU5	RICTY	782	23	82.1	470	2	P95042	MYCTU	P95042 mycobacteri
710	23	82.1	426	2	Q9A7P2	CAUCR	783	23	82.1	471	1	PBX3	HUMAN	Q9uk99 homo sapien
711	23	82.1	426	2	Q9ZCH8	RICPR	784	23	82.1	471	2	Q9H0V2	HUMAN	Q9h0v2 homo sapien
712	23	82.1	427	2	Q7Q6B4	ANOPHELES	785	23	82.1	471	2	Q86X90	HUMAN	Q86x90 homo sapien
713	23	82.1	427	2	Q75TN6	BACFI	786	23	82.1	471	2	O8FE74	ECOL6	Q8fe74 escherichia
714	23	82.1	427	2	Q75T01	BACAO	787	23	82.1	471	2	O9ROE6	9MICO	Q9rqe6 xylanimicro
715	23	82.1	427	2	Q75TX1	BACHD	788	23	82.1	474	2	O70L81	LEGPN	Q70l81 legionella
716	23	82.1	427	2	Q75TX3	BACHD	789	23	82.1	474	2	O5X4V5	LEGPA	Q5x4v5 legionella
717	23	82.1	427	2	Q75TX6	BACHD	790	23	82.1	475	2	Q4FQP5	9GAMM	Q4fqp5 psychobact
718	23	82.1	427	2	Q75TX7	BACHD	791	23	82.1	475	2	RAB31	HUMAN	Q9eqf0 homo sapien
719	23	82.1	427	2	Q75T25	BACHD	792	23	82.1	476	2	O6BUY1	DEBHA	Q6buy1 debaryomyce
720	23	82.1	427	2	Q75T27	BACHD	793	23	82.1	476	2	O5RRP1	PONPY	Q5rrp1 pongo pygma
721	23	82.1	427	2	Q75T28	BACHD	794	23	82.1	480	1	PBX3	MOUSE	Q9dc63 mus musculu
722	23	82.1	427	2	Q73V81	MYCOPACTERI	795	23	82.1	480	1	Q5UZE8	HALMA	Q5uze8 haloarcula
723	23	82.1	427	2	Q89K16	BRADYRHIZOB	796	23	82.1	480	2	Q8C710	MOUSE	Q8c710 mus musculu
724	23	82.1	427	2	Q9K5P7	BACHD	797	23	82.1	480	2	Q9D2J5	MOUSE	Q9d2j5 mus musculu
725	23	82.1	428	1	R5MB	SHEON	798	23	82.1	480	2	O8CEJ0	MOUSE	Q8cej0 mus musculu
726	23	82.1	428	2	Q7P9T3	RICKETTSIA	799	23	82.1	481	2	O873C7	NEUCR	Q873c7 neurospora
727	23	82.1	428	2	Q2GD7	RICCN	800	23	82.1	483	2	O4ZQ56	PSESY	Q4zq56 pseudomonas
728	23	82.1	429	2	Q5ZK99	GALLUS GALL	801	23	82.1	484	2	Q7Z0N8	PART	Q7z0n8 paramacium
729	23	82.1	429	2	Q7ZWA9	BRACHYDANIO	802	23	82.1	485	2	O89CL4	BRADYRHIZOB	Q89cl4 bradyrhizob
730	23	82.1	430	2	Q5RBZ1	PONPY	803	23	82.1	492	2	O4UBZ8	THEAN	Q4ubz8 theileria a
731	23	82.1	431	2	O66I16	XENTR	804	23	82.1	492	2	O76EA6	PSESG	Q76ea6 pseudomonas
732	23	82.1	431	2	Q7SYG0	ACIBE	805	23	82.1	492	2	O6I807	PSESG	Q6i807 pseudomonas
733	23	82.1	433	2	Q4UHL1	THEAN	806	23	82.1	492	2	O884Y5	PSESM	Q884y5 pseudomonas
734	23	82.1	433	2	Q64807	ARATH	807	23	82.1	495	2	Q9HFJ1	NEUCR	Q9hfj1 neurospora
735	23	82.1	434	2	Q832B6	ENTFA	808	23	82.1	501	2	Q5DF74	SCHJA	Q5df74 schistosoma
736	23	82.1	434	2	Q9DRL0	9INFA	809	23	82.1	501	2	O6KB68	9POTY	Q6kb68 sweet potat
737	23	82.1	434	2	Q9DRL1	9INFA	810	23	82.1	501	2	O599X8	9POTY	Q599x8 sweet potat
738	23	82.1	434	2	Q9DRL2	9INFA	811	23	82.1	501	2	O8B323	9POTY	Q8b323 sweet potat
739	23	82.1	434	2	Q8AVM2	XENLA	812	23	82.1	501	2	O8B325	9POTY	Q8b325 sweet potat
740	23	82.1	435	2	Q5P396	AZOSE	813	23	82.1	501	2	Q8B326	9POTY	Q8b326 sweet potat
741	23	82.1	437	2	Q4HX18	GTBZE	814	23	82.1	502	2	O8B327	9POTY	Q8b327 sweet potat
742	23	82.1	437	2	Q87H22	VIBPA	815	23	82.1	502	2	O6NGC7	CORDI	Q6ngc7 corynebacte
743	23	82.1	440	2	O86ZW3	LEPMC	816	23	82.1	502	2	O7MMH5	VIBVY	Q7mmh5 vibrio vuln
744	23	82.1	440	2	O5JTV9	HUMAN	817	23	82.1	503	1	CP7A1	RAT	P18125 rattus norv
745	23	82.1	444	2	Q5X146	RAT	818	23	82.1	503	2	O8UWR4	9INFA	Q8uwr4 influenza a
746	23	82.1	446	2	Q59G80	HUMAN	819	23	82.1	504	1	CP7A1	CRIGR	P46634 cricetus
747	23	82.1	446	2	Q87D83	XYLFT	820	23	82.1	504	2	Q9HQY0	HALISA	Q9hqy0 halobacteri
748	23	82.1	446	2	Q9PBX6	XYLFA	821	23	82.1	504	2	Q7REB1	PLAYO	Q7reb1 plasmodium
749	23	82.1	447	2	Q4KEC1	PSEF5	822	23	82.1	505	2	Q73UQ4	MYCPA	Q73uq4 mycobacteri
750	23	82.1	448	2	Q4H843	9DEIO	823	23	82.1	506	2	Q9HI92	THEAC	Q9hi92 thermoplaas
751	23	82.1	451	2	Q5JW00	HUMAN	824	23	82.1	508	2	Q4WLA0	ASPFU	Q4wla0 aspergillus
752	23	82.1	451	2	Q4NA38	9MICC	825	23	82.1	508	2	Q5FQNZ	GLUOX	Q5fqnz gluconobact
753	23	82.1	451	2	Q9A392	CAUCR	826	23	82.1	508	2	Q8DKJ2	SYNEL	Q8dkj2 synechococc
754	23	82.1	451	2	Q5ILH1	CHICK	827	23	82.1	509	2	Q4IDJ1	GIBBEZ	Q4idj1 gibberella
755	23	82.1	452	2	Q4P062	USTMA	828	23	82.1	513	2	O5CXR6	CRYPV	Q5cxr6 cryptospori
756	23	82.1	452	2	Q4P721	USTMA	829	23	82.1	513	2	O5CIH4	CRYHO	Q5cih4 corynebacte
757	23	82.1	452	2	Q8EYJ0	LEPIN	830	23	82.1	513	2	O6NGC3	CORDI	Q6ngc3 corynebacte
758	23	82.1	452	2	Q72M16	LEPIC	831	23	82.1	514	2	Q7NCS3	GLOVI	Q7ncs3 gloebacter
759	23	82.1	454	2	Q4ZLS9	PSESY	832	23	82.1	514	2	O5FHA7	EHRRG	Q5fha7 ehrlichia r
760	23	82.1	454	2	Q7SYQ3	XENLA	833	23	82.1	514	2	O5HC18	EHRRW	Q5hc18 ehrlichia r
761	23	82.1	455	2	Q87UI9	PSESM	834	23	82.1	517	2	Q5RJG8	MOUSE	Q5rjg8 mus musculu

835	23	82.1	519	2	Q5B292	EMENI	Q5b292	aspergillus	908	23	82.1	601	2	Q54IK5	DICDI	Q54ik5	dictyosteli	908
836	23	82.1	522	2	Q83B02	COXBU	Q83bq2	coxiiella bu	909	23	82.1	604	2	Q4HH51	CAMCO	Q4hh51	campylobact	909
837	23	82.1	523	2	Q95U04	DROME	Q95u04	drosophila	910	23	82.1	608	2	Q5E4C5	VIBF1	Q5e4c5	vibrio fisc	910
838	23	82.1	523	2	Q9VP67	DROME	Q9vp67	drosophila	911	23	82.1	609	2	Q8ZV0	PYRAE	Q8zvv0	pyrobaculum	911
839	23	82.1	523	2	Q9HGF6	STAAC	Q9hgf6	staphylococ	912	23	82.1	609	2	Q4THX1	TETNG	Q4thx1	tetraodon n	912
840	23	82.1	525	2	Q89R24	BRAXA	Q89r24	bradyrhizob	913	23	82.1	610	1	UVRG	ECOS7	Uvrg	eschericchia	913
841	23	82.1	527	2	Q05855	MYCTU	Q05855	mycobacteri	914	23	82.1	610	1	UVRG	ECOL6	Uvrg	eschericchia	914
842	23	82.1	530	1	LNT RHLO		Q9bhj6	rhizobium l	915	23	82.1	610	1	UVRG	ECOLI	Uvrg	eschericchia	915
843	23	82.1	531	2	Q9HIE2	THEAC	Q9hie2	thermoplasm	916	23	82.1	610	1	UVRG	SALTY	Uvrg	salmonella	916
844	23	82.1	533	2	Q4INC3	GIBZE	Q4inc4	gibberella	917	23	82.1	610	1	UVRG	VERPE	Uvrg	versinia pe	917
845	23	82.1	534	2	Q94FM4	TOBAC	Q94fm4	nicotiana t	918	23	82.1	610	1	Q5SNT7	HUMAN	Q5snt7	homo sapien	918
846	23	82.1	535	2	Q51IY3	MAGGR	Q51iy3	magnaporthe	919	23	82.1	610	2	Q57NS6	SALCH	Q57ns6	salmonella	919
847	23	82.1	536	2	Q50ME1	ENTHI	Q50me1	entamoeba h	920	23	82.1	610	2	Q5P1I0	SALPA	Q5p1i0	salmonella	920
848	23	82.1	536	2	Q87BL3	XYLFT	Q87bl3	xyliella fas	921	23	82.1	610	2	Q66BN5	YERPS	Q66bn5	yerinia ps	921
849	23	82.1	536	2	Q9PAS5	XYLFA	Q9pas5	xyliella fas	922	23	82.1	610	2	Q6D363	ERWCT	Q6d363	erwinia car	922
850	23	82.1	537	2	Q94FM5	TOBAC	Q94fm5	nicotiana t	923	23	82.1	610	2	P73400	SYNY3	P73400	synechocyst	923
851	23	82.1	539	2	Q5H2E3	XANOR	Q5hze3	xanthomonas	924	23	82.1	611	1	UVRG	PASNU	Uvrg	pasteurella	924
852	23	82.1	543	2	Q5A1C8	CANAL	Q5alc8	xandida alb	925	23	82.1	611	2	Q69506	MYCLE	Q69506	mycobacteri	925
853	23	82.1	543	2	Q4A1C8	MYCTU	Q4a1c8	mycobacteri	926	23	82.1	611	2	Q53666	MYCTU	Q53666	mycobacteri	926
854	23	82.1	544	2	Q9CCH3	MYCLE	Q9cch3	mycobacteri	927	23	82.1	611	2	Q7U2H0	MYCBO	Q7u2h0	mycobacteri	927
855	23	82.1	545	2	Q6RQZ4	DEBHA	Q6bqz4	debaromyce	928	23	82.1	613	2	P72613	SYNY3	P72613	synechocyst	928
856	23	82.1	548	2	Q71MJ3	CITPS	Q71mj3	citrus sine	929	23	82.1	619	2	Q4Q268	LEIMA	Q4q268	leishmania	929
857	23	82.1	548	2	Q8RVR2	CITPA	Q8rvr2	citrus para	930	23	82.1	620	2	Q9HX88	PSEAE	Q9hx88	pseudomonas	930
858	23	82.1	555	2	Q6CIU4	KLULA	Q6ciu4	kluyveromyc	931	23	82.1	621	2	Q6Y261	ORISA	Q6y261	oryza sativ	931
859	23	82.1	555	2	Q9FYU6	9ROSI	Q9fyu6	citrus juno	932	23	82.1	626	1	UVRG	SYNY3	Uvrg	synechocyst	932
860	23	82.1	557	2	Q4L5X8	STAHJ	Q4l5x8	staphylococ	933	23	82.1	627	1	UVRG	ANASP	Uvrg	anabaena sp	933
861	23	82.1	557	2	Q6G9T8	STNAA	Q6g9t8	staphylococ	934	23	82.1	631	2	Q4HMT1	CAMLA	Q4hmt1	campylobact	934
862	23	82.1	557	2	Q6GHG0	STAAR	Q6ghg0	staphylococ	935	23	82.1	632	1	STC CLOBE		Stc	clostridium	935
863	23	82.1	557	2	Q7A115	STAAN	Q7a115	staphylococ	936	23	82.1	632	2	Q7Z5G9	HUMAN	Q7z5g9	homo sapien	936
864	23	82.1	557	2	Q7A5X6	STAAN	Q7a5x6	staphylococ	937	23	82.1	632	2	Q89KU2	BRAJA	Q89ku2	bradyrhizob	937
865	23	82.1	557	2	Q99UJ7	STAAM	Q99uj7	staphylococ	938	23	82.1	633	2	Q5NYD2	AZOSE	Q5nyd2	azocarcus sp	938
866	23	82.1	559	2	Q51LU9	MAGGR	Q51lu9	magnaporthe	939	23	82.1	634	2	Q8H150	ARATH	Q8h150	arabidopsis	939
867	23	82.1	559	2	Q8H144	ARATH	Q8h144	arabidopsis	940	23	82.1	634	2	Q9FN91	ARATH	Q9fn91	arabidopsis	940
868	23	82.1	562	2	Q7Y8D7	9PLEO	Q7y8d7	cochilobolu	941	23	82.1	634	2	Q6APJ9	DRSPS	Q6apj9	desulfotale	941
869	23	82.1	562	2	Q7JK62	CAEEL	Q7jk62	caenorhabdi	942	23	82.1	640	2	Q7N531	PHOLL	Q7n531	photorhabdu	942
870	23	82.1	565	2	Q6FT63	CANGA	Q6ft63	candida gla	943	23	82.1	641	2	Q5B2V6	EMENI	Q5b2v6	aspergillus	943
871	23	82.1	566	1	AMY STRLM		P03270	streptomyce	944	23	82.1	645	2	Q5N031	SYNP6	Q5n031	synechococ	944
872	23	82.1	566	1	AMY STRLM		P09794	streptomyce	945	23	82.1	645	2	Q4LIX2	9BURL	Q4lix2	burkholderi	945
873	23	82.1	568	1	UVRG AQUAE		Q67887	aquifex aeo	946	23	82.1	645	2	Q5FNM9	GLUOX	Q5fnm9	gluconobact	946
874	23	82.1	568	1	Q5HPR6	STAEQ	Q5hpr6	staphylococ	947	23	82.1	650	2	Q5LNT9	SILPO	Q5lnt9	silicibacte	947
875	23	82.1	572	2	Q8GST0	STAEP	Q8gst0	staphylococ	948	23	82.1	651	2	Q5BBC9	EMENI	Q5bbc9	aspergillus	948
876	23	82.1	573	2	Q9USN2	MYZPE	Q9usn2	myzus persi	949	23	82.1	652	2	Q9VS56	DROME	Q9vs56	drosophila	949
877	23	82.1	573	2	Q70270	ANOGA	Q7q270	anopheles g	950	23	82.1	654	2	Q4WJ23	ASPFU	Q4wj23	aspergillus	950
878	23	82.1	574	2	Q6Q7P2	ECOLI	Q6q7p2	eschericchia	951	23	82.1	654	2	Q712J2	RAT	Q712j2	rattus norv	951
879	23	82.1	574	2	Q573I8	9BACT	Q573i8	uncultured	952	23	82.1	656	2	Q7V1I9	PROMP	Q7v1i9	prochloroco	952
880	23	82.1	575	1	ENV SMRVH		P21412	squirrei mo	953	23	82.1	661	2	Q7VX63	BORPE	Q7vx63	bordetella	953
881	23	82.1	575	1	IUCA ECOLI		Q47316	eschericchia	954	23	82.1	661	2	Q7W829	BORBA	Q7w829	bordetella	954
882	23	82.1	579	1	UVRG METTH		Q26541	methanobact	955	23	82.1	661	2	Q7W1H6	BORBA	Q7w1h6	bordetella	955
883	23	82.1	580	2	Q4WEV3	ASPFU	Q4wbv3	aspergillus	956	23	82.1	666	2	Q5AVX6	EMENI	Q5avx6	aspergillus	956
884	23	82.1	581	2	Q46219	9DIPT	Q46219	culicoides	957	23	82.1	666	2	Q5AVX6	EMENI	Q5avx6	aspergillus	957
885	23	82.1	588	2	Q83R46	SHIFL	Q83r46	shigella fi	958	23	82.1	666	2	Q5AVX6	EMENI	Q5avx6	aspergillus	958
886	23	82.1	590	2	Q610B6	CABER	Q610b6	caenorhabdi	959	23	82.1	666	2	Q5AVX6	EMENI	Q5avx6	aspergillus	959
887	23	82.1	591	2	Q4WFN9	ASPFU	Q4wfn9	aspergillus	960	23	82.1	667	2	Q5Z230	NOCFA	Q5z230	nocardia fa	960
888	23	82.1	592	2	Q92Y36	RHIME	Q92y36	rhizobium m	961	23	82.1	669	1	Q7V7L7	PROMM	Q7v7l7	prochloroco	961
889	23	82.1	593	2	Q6KD48	ECOLI	Q6kd48	eschericchia	962	23	82.1	669	1	UVRB	XYLFA	Uvrb	xyliella fas	962
890	23	82.1	593	2	Q93PE6	SHIBO	Q93pe6	shigella bo	963	23	82.1	669	1	UVRB	XYLFT	Uvrb	xyliella fas	963
891	23	82.1	593	2	Q9XCH4	SHIFL	Q9xch4	shigella fi	964	23	82.1	672	2	Q5WNM3	CABER	Q5wnm3	caenorhabdi	964
892	23	82.1	594	1	UL93 HCMVA		P16799	human cytom	965	23	82.1	673	1	UVRB	XANAC	Uvrb	xanthomonas	965
893	23	82.1	594	2	Q4KAL6	PSEF5	Q4kal6	pseudomonas	966	23	82.1	673	1	UVRB	XANCP	Uvrb	xanthomonas	966
894	23	82.1	594	2	Q6RXE6	HCMV	Q6rx66	human cytom	967	23	82.1	673	2	Q9RBK2	XANCP	Q9rbk2	xanthomonas	967
895	23	82.1	594	2	Q80MY0	HCMV	Q80my0	human cytom	968	23	82.1	674	2	Q62140	CABEL	Q62140	caenorhabdi	968
896	23	82.1	594	2	Q6SW50	HCMV	Q6sw50	human cytom	969	23	82.1	677	2	Q81Q81	DROME	Q81q81	drosophila	969
897	23	82.1	595	2	Q8FDV6	ECOL6	Q8fdv6	eschericchia	970	23	82.1	677	2	Q7ZBN0	DESVH	Q7zbn0	desulfovibr	970
898	23	82.1	597	2	Q5R619	PONPY	Q5r619	pongo pygma	971	23	82.1	678	2	Q4UXG7	XANCP	Q4uxg7	xanthomonas	971
899	23	82.1	597	2	Q4HIV4	CAMLA	Q4hiv4	campylobact	972	23	82.1	678	2	Q8P6P5	XANCP	Q8p6p5	xanthomonas	972
900	23	82.1	598	2	Q55FN1	DICDI	Q55fn1	dictyosteli	973	23	82.1	678	2	Q8PHZ9	XANCP	Q8phz9	xanthomonas	973
901	23	82.1	598	2	Q8GHE7	SALTY	Q8ghe7	salmonella	974	23	82.1	678	2	Q62137	CABEL	Q62137	caenorhabdi	974
902	23	82.1	598	2	Q939L2	CITPR	Q939l2	citrobacter	975	23	82.1	679	2	O6K1G9	MOUSE	O6k1g9	mus musculu	975
903	23	82.1	598	2	Q57J50	SALCH	Q57j50	salmonella	976	23	82.1	687	2	Q7VCW1	PROMA	Q7vcw1	prochloroco	976
904	23	82.1	598	2	Q5XDP2	ENTAM	Q5xdp2	enterobacte	977	23	82.1	688	2	Q5GXX4	XANOR	Q5gxx4	xanthomonas	977
905	23	82.1	598	2	Q5PMU8	SALPA	Q5pmu8	salmonella	978	23	82.1	689	1	ACOX1	YARLI	Acox1	yarrowia li	978
906	23	82.1	598	2	Q8Z3N6	SALTY	Q8z3n6	salmonella	979	23	82.1	689	1	ACOX1	YARLI	Acox1	yarrowia li	979
907	23	82.1	600	2	Q51EP9	ENTHI	Q51ep9	entamoeba h	980	23	82.1	689	1	ACOX1	YARLI	Acox1	yarrowia li	980

```
981 23 82.1 691 2 Q81Q80 drosophila
982 23 82.1 693 2 Q4P2J2 ustilago ma
983 23 82.1 693 2 Q87D39 xylella fas
984 23 82.1 693 2 Q9PC32 xylella fas
985 23 82.1 694 2 Q4UDG8 theileria a
986 23 82.1 695 1 ELF3 ARATH
987 23 82.1 697 2 Q4N630 arabidopsis
988 23 82.1 697 2 Q4N630 theileria p
989 23 82.1 700 1 Q8QV88 eyach virus
990 23 82.1 702 1 AC0X33 YARLI
991 23 82.1 705 2 Q4SKH2 tetraodon n
992 23 82.1 705 2 Q9Y8D5 cochiobolu
993 23 82.1 708 2 Q7QDU7 anopheles g
994 23 82.1 718 2 Q5JVZ5 homo sapien
995 23 82.1 718 2 Q5GMG3 mus musculu
996 23 82.1 719 1 AC0X_FICPA
997 23 82.1 720 1 ELMO2 HUMAN
998 23 82.1 720 2 Q5RCC1 PONPY
999 23 82.1 723 2 Q66IU1 XENTR
1000 23 82.1 725 2 Q8DUC4 torque teno
Q6G471 bartonella
```

ALIGNMENTS

```
RESULT 1
Q6MGX7_BDEBA
ID Q6MGX7_BDEBA PRELIMINARY; PRT; 232 AA.
AC Q6MGX7
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Bd3792;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
EX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goessmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective";
RL Science 303:689-692(2004).
DR EMBL; BX842656; CAE81152.1; -; Genomic_DNA.
DR InterPro; IPR003730; DUF152.
DR Pfam; PF02578; DUF152; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 232 AA; 25776 MW; 80A2229E03ED2677 CRC64;
```

```
Query Match 96.4%; Score 27; DB 2; Length 232;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 VRDQIXXL 9
Db 148 VRDQILSSL 156
```

```
RESULT 2
Q7QAB5_ANOGA
ID Q7QAB5_ANOGA PRELIMINARY; PRT; 455 AA.
AC Q7QAB5;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP0000013373 (Fragment).
GN ORFNames=ENSANG0000010884;
OS Anopheles gambiae str. PEST.
```

```
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008898; EAA09186.2; -; Genomic_DNA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0016071; P:RNA metabolism; IEA.
DR SMART; SM00025; Pumilio; 4.
DR PROSITE; PS50303; PUM_HD; 1.
FT NON_TER 1
FT NON_TER 455
SQ SEQUENCE 455 AA; 51607 MW; 2D6DD0915F9AC2D1 CRC64;
```

```
Query Match 96.4%; Score 27; DB 2; Length 455;
Best Local Similarity 55.6%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 VRDQIXXL 9
Db 110 IRDQIATSL 118
```

```
RESULT 3
Q9ZSH5_ARATH
ID Q9ZSH5_ARATH PRELIMINARY; PRT; 543 AA.
AC Q9ZSH5;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE T9E19.1 protein.
GN Name=T9E19.1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA WashU;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Stromatt C., Johnson D., Le T.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104920; AAC72877.1; -; Genomic_DNA.
DR PIR; T02022; T02022.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase C48.
DR Pfam; PF02902; Peptidase C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 543 AA; 60444 MW; 8D2438A5F022DBF1 CRC64;
```

```
Query Match 96.4%; Score 27; DB 2; Length 543;
```

Best Local Similarity 66.7%; Pred. No. 5.2e+02; Mismatches 0; Indels 3; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 69 VRDQIRAAAL 77

RESULT 4

ID Q9M180 ARATH PRELIMINARY; PRT; 1113 AA.
AC Q9M180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein T5C2_90.
GN Name=T5C2_90;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138664; CAB81580.1; -; Genomic_DNA.
DR FIR; T47381; T47381.
KW Hypothetical protein.
SQ SEQUENCE 1113 AA; 123579 MW; B826A0588B97DDBE CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1113;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 488 VRDQIRAAAL 496

RESULT 5

ID Q9LK04 ARATH PRELIMINARY; PRT; 1119 AA.
AC Q9LK04;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gb|AAC24188.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE-20363099; PubMed-10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000382; BAB02940.1; -; Genomic DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS0600; ULP_PROTEASE; 1.
SQ SEQUENCE 1119 AA; 125821 MW; CD3483A918B0895 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1119;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 488 VRDQIRAAAL 496

RESULT 6

ID Q9S9Z6 ARATH PRELIMINARY; PRT; 1132 AA.
AC Q9S9Z6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE F21H2.5.
GN Name=F21H2.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C.,
RA Rowley D., Shinn P., Walker M., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Theologis;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007894; AAD46015.1; -; Genomic_DNA.
DR GO; GO:0008234; P:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS0600; ULP_PROTEASE; 1.
SQ SEQUENCE 1132 AA; 126285 MW; 7F04A71FE0BF6A2E CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1132;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
| | | | |
Db 468 VRDQIRAAAL 476

RESULT 7

ID Q9C6N9 ARATH PRELIMINARY; PRT; 1198 AA.
AC Q9C6N9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F28L5.2.
GN Name=F28L5.2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Fraser C.M.;
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079280; AAG50582.1; -; Genomic_DNA.
DR PIR; E86402; E86402.
DR GO; GO:0008234; P:cytosine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1198 AA; 133579 MW; D608AE577CE21339 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1198;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 486 VRDQIRATL 494

RESULT 8
Q9C607 ARATH PRELIMINARY; PRT; 1201 AA.
ID Q9C607 ARATH PRELIMINARY; PRT; 1201 AA.
AC Q9C607;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F14G11.4.
GN Name=F14G11.4;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084221; AAG50521.1; -; Genomic_DNA.
DR PIR; F86386; F86386.
DR GO; GO:0008234; P:cytosine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1201 AA; 133995 MW; D10F98847289CFB7 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1201;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 455 VRDQIRAL 463

RESULT 9
Q9ZU21 ARATH PRELIMINARY; PRT; 1305 AA.
ID Q9ZU21 ARATH PRELIMINARY; PRT; 1305 AA.
AC Q9ZU21;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F5F19.8 protein.
GN Name=F5F19.8;
OS Arabidopsis thaliana (Mouse-ear cress).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Vyotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Lenz C., Liu S.,
RA Li J., Kremenetskaia I., Luros J., Gonzalez A., Altafi H., Araujo R.,
RA Buehler B., Conway A.B., Conn L., Dunn P., Hansen N., Huizar L.,
RA Kim C., Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006216; AAD12676.1; -; Genomic_DNA.
DR PIR; H96559; H96559.
DR GO; GO:0008234; P:cytosine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1305 AA; 146079 MW; D3C036AEE2EFD6C CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1305;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 491 VRDQIRAL 499

RESULT 10
Q9LTG8 ARATH PRELIMINARY; PRT; 1444 AA.
ID Q9LTG8 ARATH PRELIMINARY; PRT; 1444 AA.
AC Q9LTG8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similarity to mutator-like transposase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025605; BAA98062.1; -; Genomic_DNA.
DR GO; GO:0008234; P:cytosine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS50600; ULP_PROTEASE; 1.
SQ SEQUENCE 1444 AA; 161024 MW; AF0CB23AAAC4EB8B CRC64;

Query Match 96.4%; Score 27; DB 2; Length 1444;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 488 VRDQIRAL 496

RESULT 11
Q9SFY0 ARATH PRELIMINARY; PRT; 1444 AA.
ID Q9SFY0 ARATH PRELIMINARY; PRT; 1444 AA.
AC Q9SFY0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T22C5.24.

```

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.R.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN EMBL; AC012375; AAF24952.1; -; Genomic DNA.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS06000; ULP_PROTEASE; 1.
SQ SEQUENCE 1444 AA; 161065 MW; C0563453E81C6498 CRC64;
Query Match 96.4%; Score 27; DB 2; Length 1444;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
DB 488 VRDQIRATL 496
RESULT 12
ID QSL11_1 ARATH PRELIMINARY; PRT; 1472 AA.
AC QSL11;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein At2G05560;
GN Name=At2G05560;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.B., Mason T.H., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.W., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006220; AAD24651.1; -; Genomic_DNA.
DR PIR; A84470; A84470.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS06000; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1472 AA; 165163 MW; FFE08B8AB3AE1787A CRC64;
Query Match 96.4%; Score 27; DB 2; Length 1472;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
DB 403 VRDQIRAAAL 411
RESULT 13
ID O80975 ARATH PRELIMINARY; PRT; 1756 AA.
AC O80975;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein At2G14770.
GN Name=At2G14770;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004705; AAC24188.1; -; Genomic_DNA.
DR PIR; T02599; T02599.
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003653; Peptidase_C48.
DR Pfam; PF02902; Peptidase_C48; 1.
DR PROSITE; PS06000; ULP_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 1756 AA; 197367 MW; 68FC3B5FCC282775 CRC64;
Query Match 96.4%; Score 27; DB 2; Length 1756;
Best Local Similarity 66.7%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 VRDQIXXXL 9
DB 488 VRDQIRAAAL 496
RESULT 14
ID O65247 ARATH PRELIMINARY; PRT; 211 AA.
AC O65247;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE F2IE10.9 protein.
GN Name=F2IE10.9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RA WashU.
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Davidson S., Rohlfing T., David M., O'Brian D.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF058914; AAC13589.1; -; Genomic_DNA.
DR PIR; T01194; T01194.
DR HSSP; P11746; 1NMN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF00319; SRF-TF; 1.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS0066; MADS_BOX 2; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 211 AA; 24251 MW; FBE93738384E60D CRC64;

Query Match 92.9%; Score 26; DB 2; Length 211;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 187 VRDQIQITL 195
|||||
187 VRDQIQITL 195

RESULT 15
Q7XJK7 ARATH PRELIMINARY; PRT; 218 AA.
AC Q7XJK7
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE MADS-box protein AGU35.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC TISSUE-Slignics;
RA Parenicova L., de Folter S., Kieffer M., Horner D.S., Favalli C.,
RA Buescher J., Cook H.E., Ingram R.M., Kater M.M., Davies B.,
RA Angenot G.C., Colombo L.;
RT Molecular and Phylogenetic Analyses of the Complete MADS-Box
Transcription Factor Family in Arabidopsis: New Openings to the MADS
World.;
RT Plant Cell 0:0-0(2003).
RL EMBL; AY141246; AA52810.1; -; mRNA.
DR HSSP; P11831; 1HBX.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF00319; SRF-TF; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS0066; MADS_BOX 2; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 218 AA; 25123 MW; 3D8DF0E56CDBCD4 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 218;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 194 VRDQIQITL 202
|||||
194 VRDQIQITL 202

RESULT 16
Q5WH83 BACSK PRELIMINARY; PRT; 272 AA.
AC Q5WH83;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Phosphomethylpyrimidine kinase (EC 2.7.4.7).
GN OrderedLocustNames=ABC1737;
OS Bacillus clausii (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=66692;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN-KSM-K16;
RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA Kawai S., Ito S., Horikoshi K.;
RT "The complete genome sequence of the alkaliphilic Bacillus clausii
KSM-K16";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006627; BAD64272.1; -; Genomic_DNA.
DR GO; GO:0008972; F:phosphomethylpyrimidine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009228; P:thiamin biosynthesis; IEA.
DR InterPro; IPR004399; HMP-P_kinase.
DR InterPro; IPR011611; PfkB_region.
DR Pfam; PF00294; PfkB; 1.
DR TIGRFAMs; TIGR00097; HMP-P_kinase; 1.
KW Complete proteome; Kinase; Transferase.
SQ SEQUENCE 272 AA; 28197 MW; A81DA7A12A298AF7 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 272;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 60 VRDQIEAVL 68
|||||
60 VRDQIEAVL 68

RESULT 17
Q7NKK0 GLOVI PRELIMINARY; PRT; 334 AA.
AC Q7NKK0;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Gll1477 protein.
GN OrderedLocustNames=gll1477;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP


```
RC STRAINS=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003);
DR EMBL; BA000045; BAC89418.1; -; Genomic_DNA.
DR InterPro; IPR007296; DUF403_bac.
DR Pfam; PF04168; DUF403; 1.
KW Complete proteome.
SQ SEQUENCE 334 AA; 38221 MW; 44519A598FBE1920 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 334;
Best Local Similarity 55.6%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 100 IRDQISSEL 108

RESULT 18
QB2110 SALT1 PRELIMINARY; PRT; 357 AA.
AC QB2110; Q7C426;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-SEP-2005 (TrEMBLrel. 29, Last annotation update)
DE Protein kinase.
GN Ordered locus names=STY4823, t4520;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churser C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627283; CAD06945.1; -; Genomic_DNA.
DR EMBL; AE016849; AA071965.1; -; Genomic_DNA.
DR HSSP; P71584; 1MRU.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004572; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
```

```
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Transferase.
SQ SEQUENCE 357 AA; 40433 MW; 5C234A0F141516D6 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 255 VRDQIAKYL 263

RESULT 19
QBMYV0 DROME PRELIMINARY; PRT; 486 AA.
AC QBMYV0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RH42429p.
GN ORFNames=CG7433;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkelley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorset V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
DR EMBL; AY113591; AAM29596.1; -; mRNA.
DR HSSP; P80147; 1OHV.
DR Ensembl; CG7433; Drosophila melanogaster.
DR FlyBase; FBgn0036927; CG7433.
DR GO; GO:0003867; F:4-aminobutyrate transaminase activity; IEA.
DR GO; GO:0030170; F:pyridoxal phosphate binding; IEA.
DR GO; GO:0009448; P:gamma-aminobutyric acid metabolism; IEA.
DR InterPro; IPR005814; Aminotrans_3.
DR InterPro; IPR004631; GABAtrns_euk.
DR PANTHER; PTHR11986; Aminotrans_3; 1.
DR Pfam; PF00202; Aminotran_3; 1.
DR TIGRFAMs; TIGR00699; GABAtrns_euk; 1.
KW Pyridoxal phosphate.
SQ SEQUENCE 486 AA; 54530 MW; 7A1C8F1140388D78 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
:|||||
Db 436 VRDQIIGAL 444

RESULT 20
QBVM68 DROME PRELIMINARY; PRT; 486 AA.
AC QBVM68;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG7433-PA, isoform A (CG7433-pb, isoform b).
GN Name=CG7433; ORFNames=CG7433;
OS Drosophila melanogaster (Fruit fly).
```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN [1]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Gordon R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle B.J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celnik S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacle B.J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby H.A., Munkall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celnik S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]

RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celnik S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacle B.J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]

RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- INTERACTION:
CC OSVWR8-1:CG6659; NbExp=1; IntAct=EBI-86087; EBI-281215;
CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC EMBL; AB003515; AAF49081.1; -; Genomic_DNA.
DR HSP; P80147; 1OHV.
DR IntAct; Q9VM68; -.
DR Ensembl; CG7433; Drosophila melanogaster.
DR FlyBase; FBgn0036927; CG7433
DR GO; GO:0003867; F:4-aminobutyrate transaminase activity; IEA.
DR GO; GO:0030170; F:Pyridoxal phosphate binding; IEA.
DR GO; GO:0009448; P:Gamma-aminobutyric acid metabolism; IEA.
DR InterPro; IPR005814; Aminotrans_3.
DR InterPro; IPR004631; GABATrans_euk.
DR PANTHER; PTHR11986; Aminotrans_3; 1.
DR Pfam; PF00202; Aminotran_3; 1.
DR TIGRFAMs; TIGR00699; GABATrans_euk; 1.
KW Pyridoxal phosphate.
SQ SEQUENCE 486 AA; 54584 MW; 405BF910A543A51B CRC64;

Query Match 92.98; Score 26; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXX 9
Db 436 VRDQIIGAL 444

RESULT 21

ID Q4PCC1_USTMA PRELIMINARY; PRT; 516 AA.
AC Q4PCC1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM02242.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Bouhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,

```

RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Govette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hubby E., Iliev I.,
RA Jaffe D., Jones C., Kanat M., Kanat A., Kamyssekis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lema D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-Toh K., Liu X., Lorkitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbutt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Maraballa R., Maru K., Matthews C., Maudeli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalov A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Moses J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okawa O., O'leary S., Omotosho B.,
RA O'Neill K., Oman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Teefaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACPO1000079; BAK93364.1; -: Genomic_DNA.
DR Hypothetical protein.
KW SEQUENCE 516 AA; 54638 MW; 2B80AF2853AE4B65 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 516;
Best Local Similarity 66.7%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 382 VRDQITNHL 390

RESULT 22
Q4Q7Y6 LEIMA
ID Q4Q7Y6 LEIMA PRELIMINARY; PRT; 594 AA.
AC Q4Q7Y6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Gamma-tubulin complex subunit, putative.
GN ORFNames=LmjF28.2760;
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CTO05266; CAJ05734.1; -: Genomic DNA.
KW SEQUENCE 594 AA; 68118 MW; 566DB8C1434C721B CRC64;

Query Match 92.9%; Score 26; DB 2; Length 594;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9

```

```

DB 320 VRDQVQSAL 328

RESULT 23
Q89S89 BRAJA
ID Q89S89 BRAJA PRELIMINARY; PRT; 642 AA.
AC Q89S89;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B112516 protein.
GN OrderedLocusNames=b112516;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC47781.1; -: Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 642 AA; 68447 MW; A3F527A4F4323CFF CRC64;

Query Match 92.9%; Score 26; DB 2; Length 642;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 416 VRDQIEADL 424

RESULT 24
Q5LX00 SILPO
ID Q5LX00 SILPO PRELIMINARY; PRT; 743 AA.
AC Q5LX00;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SPO0204;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran W.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pal G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment.";
RL Nature 432:910-913(2004).
DR EMBL; CP000031; RAV93530.1; -: Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 743 AA; 81225 MW; 00B63B9D70C629BE CRC64;

Query Match 92.9%; Score 26; DB 2; Length 743;

```

Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
|||||
Db 731 VRDQIDSL 739

RESULT 25

ID Q7UYC9_RHOBA PRELIMINARY; PRT; 896 AA.
AC Q7UYC9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable chromosome segregation protein.
GN OrderedLocusNames=RE714;
OS Rhodospirillum rubrum
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borsym K., Heitmann K., Rabus R.,
RA Schleiner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294134; CAD71745.1; -; Genomic_DNA.
DR GO; GO:0006118; P:heme binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR011282; Cytochrome_c_r.
DR InterPro; IPR011444; DUF1549.
DR InterPro; IPR011429; PSC3.
DR Pfam; PF07635; PSCytl; 1.
DR Pfam; PF07583; PSCytl; 1.
DR Pfam; PF07587; PSD1; 1.
KW Complete proteome.
SQ SEQUENCE 896 AA; 101882 MW; 6C5627C8EEA742CE CRC64;

Query Match 92.9%; Score 26; DB 2; Length 896;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
|||||
Db 306 VRDQIAGDL 314

RESULT 26

ID Q65353_NPVAC PRELIMINARY; PRT; 1084 AA.
AC Q65353;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF B.
OS Autographa californica nuclear polyhedrosis virus (ACNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=mutant FP-D;
RX MEDLINE=90258898; PubMed=1692964;
RA Friesen P.D., Nissen M.S.;
RT "Gene organization and transcription of TED, a lepidopteran
RT retrotransposon integrated within the baculovirus genome.";
RL Mol. Cell. Biol. 10:3067-3077(1990).
DR EMBL; M32662; AAA92249.1; -; Genomic_DNA.

DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT_1; 1.
SQ SEQUENCE 1084 AA; 125586 MW; 77E8D0063F83F296 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXL 9
|||||
Db 154 VRDQITKML 162

RESULT 27

ID Q4P3H5_USTMA PRELIMINARY; PRT; 1105 AA.
AC Q4P3H5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05338.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=437631;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Arnbruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavsky L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Fitzgerald M., Foley K., Gage D., Ferreira P., Fischer H.,
RA Frickson J., Farina A., Faro S., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Hueby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kanvaseelis M., Karlsson E.,
RA Kells C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Minova T., Mikkelson T., Miengva V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,

RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of *Ustilago maydis*";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AACP01000192; EAK86587.1; -; Genomic_DNA.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR003163; Yeast_DNA_bd.
DR Pfam: PF00023; ANK; 3.
DR Pfam: PF02292; APSES; 1.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 2.
DR PROSITE: PS50297; ANK REP REGION; 1.
DR PROSITE: PS50088; ANK_REPEAT; 1.
KW ANK repeat; DNA-binding; Hypothetical protein; Repeat.
SQ SEQUENCE 1105 AA; 115924 MW; CDCBA6DAABE398DF CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1105;
Best Local Similarity 55.6%; Pred. No. 2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
DB 814 VRDQLSSTL 822

RESULT 28
Q411R0 GIBZE
ID Q411R0 GIBZE PRELIMINARY; PRT; 1191 AA.
AC Q411R0
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG08848.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gaidyna S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Huime W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kelle C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Mihova T., Miengva V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramaamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer S., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Testaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainbek L., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: AACM01000358; EAA72022.1; -; Genomic_DNA.
DR DR
KW Hypothetical protein.
SQ SEQUENCE 1191 AA; 135430 MW; A09A78310DC82C38 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1191;
Best Local Similarity 55.6%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
DB 804 IRDQILSAL 812

RESULT 29

ID Q9M0W7 ARATH PRELIMINARY; PRT; 1312 AA.
AC Q9M0W7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein At4G05280.
GN Name=At4G05280;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL161503; CAB81070.1; -; Genomic_DNA.
DR PIR: D85066; D85066.
DR GO: GO:0008234; F:cysteine-type peptidase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR003653; Peptidase_C48.
DR Pfam: PF02902; Peptidase_C48; 1.
DR PROSITE: PS50600; ULP_PROTEASE; 1.
DR Hypothetical protein.
SQ SEQUENCE 1312 AA; 147144 MW; 56269EEEE635C128 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1312;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
| | | | |
DB 434 VRDQIRAVL 442

RESULT 30

Q61Z65 CAEBR PRELIMINARY; PRT; 1407 AA.
ID Q61Z65 CAEBR PRELIMINARY; PRT; 1407 AA.
AC Q61Z65
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG03251 (Fragment).
GN Name=CBG03251;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: CAAC01000012; CAE59791.1; -; Genomic_DNA.
DR GO: GO:0004197; F:cysteine-type endopeptidase activity; IEA.

DR GO: 0004221; P: ubiquitin thiolesterase activity; IEA.
 DR GO: 0006512; P: ubiquitin cycle; IEA.
 DR GO: 0006511; P: ubiquitin-dependent protein catabolism; IEA.
 DR InterPro: IPR001394; Peptidase_C19.
 DR Pfam: PF00443; UCH_1.
 DR PROSITE: PS00972; UCH_2_1; 1.
 DR PROSITE: PS00973; UCH_2_2; 1.
 DR PROSITE: PS0235; UCH_2_3; 1.
 KW Hypothetical protein.
 FT NON_TER 1 1407
 SQ SEQUENCE 1407 AA; 164607 MW; 139EC18FC9677E4A CRC64;

Query Match 92.9%; Score 26; DB 2; Length 1407;
 Best Local Similarity 66.7%; Pred. No. 2.6e+03;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 135 VRDQIQSVL 143

RESULT 31

Q51WD5 MAGGR
 ID Q51WD5 MAGGR PRELIMINARY; PRT; 2240 AA.
 AC Q51WD5
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG03850.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
 OX NCBI_TaxID=242507;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STPAIN70-15;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
 RA Bayul T., Blitshetev B., Bloom T., Blye J., Boguslavskiy L.,
 RA Borowsky M., Boukhgaltier B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Geavin G., Gnerre S.,
 RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hegopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
 RA Jaffe D., Jones C., Kanai M., Kanat A., Kamveselis M., Karlsson E.,
 RA Kells C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
 RA Lana D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokytsang T., Lokytsang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
 RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
 RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Menga V., Moru K.,
 RA Moraes J., Mulrain L., Munson G., Navlor J., Neves C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
 RA O'Neill K., O'man S., Parker S., Perrin D., Phunkhang P., Piquani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
 RA Sheridan J., Sharpa N., Shi J., Smirnov S., Smith C., Senguez C.,
 RA Spencer B., Staiker J., Stange-thomann N., Stavropoulos S.,
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
 RA Tenzing P., Testaye S., Theodore J., Thoulutang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
 RA Venkateraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,

RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Lander E.;
 RT "The genome sequence of Magnaporthe grisea.";
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AACU01001031; EAA50091.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 2240 AA; 255489 MW; 5E00710960191B59 CRC64;

Query Match 92.9%; Score 26; DB 2; Length 2240;
 Best Local Similarity 66.7%; Pred. No. 4.2e+03;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 |||||
 Db 617 VRDQIKAKL 625

RESULT 32

Q76YV2_9CAUD
 ID Q76YV2_9CAUD PRELIMINARY; PRT; 89 AA.
 AC Q76YV2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AehLORFI29C;
 OS Bacteriophage Aeh1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=227470;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Petrov V., Nolan J., Bertrand C., Letarov A.V., Krusch H.M.,
 RA Karam J.D.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Petrov V., Nolan J., Karam J.D.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY266303; AAQ17794.1; -; Genomic_DNA.
 DR GO: 0003677; F: DNA binding; IEA.
 DR GO: 0006355; P: regulation of transcription, DNA-dependent; IEA.
 KW Hypothetical protein.
 SQ SEQUENCE 89 AA; 10144 MW; 572457907A444EA8 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 89;
 Best Local Similarity 44.4%; Pred. No. 2.5e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
 :|||:
 Db 58 IRDQVAAKL 66

RESULT 33

Q6RXD6_HCMV
 ID Q6RXD6_HCMV PRELIMINARY; PRT; 102 AA.
 AC Q6RXD6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

```
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UL101.
GN Name=UL101;
OS Human cytomegalovirus (HHV-5) (Human herpesvirus 5).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OC NCBI_TaxID=10359;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Toledo;
RA Brondke H., Schmitz B., Shenk T., Doerfler W.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY486477; AAS48987.1; -; Genomic DNA.
SQ SEQUENCE 102 AA; 10795 MW; 0CFFF4C288E02B57 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 102;
Best Local Similarity 55.6%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
   ||||:
DB 60 VRDQLSSQL 68

RESULT 34
UL101_HCMVA STANDARD; PRT; 115 AA.
ID UL101_HCMVA
AC P16826;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein UL101.
GN Name=UL101;
OS Human cytomegalovirus (strain AD169) (HHV-5) (Human herpesvirus 5).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OC NCBI_TaxID=10360;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Hornell T., Hutchinson C.A. III, Kuzarides T., Martignetti J.A.,
RA Freddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: X17403; CAA35337.1; -; Genomic_DNA.
DR FIR; S09866; S09866.
KW Hypothetical protein.
SQ SEQUENCE 115 AA; 12184 MW; 5D3BDC55D48D904F CRC64;

Query Match 89.3%; Score 25; DB 1; Length 115;
Best Local Similarity 55.6%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
   ||||:
DB 73 VRDQLSSQL 81

RESULT 35
Q58J60_9ACTO PRELIMINARY; PRT; 130 AA.
ID Q58J60_9ACTO
AC Q58J60_9ACTO
DT 10-MAY-2005 (TrEMBLrel. 30, Created)

Query Match 89.3%; Score 25; DB 2; Length 139;
Best Local Similarity 55.6%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Putative regulatory protein.
OS Streptomyces noursei ATCC 11455.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=316284;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 11455;
RX PubMed=15990252; DOI=10.1016/j.femsle.2005.05.052;
RA Volokhan O., Sletta H., Sekurova O.N., Ellingsen T.E., Zotchev S.B.;
RT "An unexpected role for the putative 4'-phosphopantetheinyl
RT transferase-encoding gene nysf in the regulation of nystatin
RT biosynthesis in Streptomyces noursei ATCC 11455.";
RL FEMS Microbiol. Lett. 249:57-64(2005).
DR EMBL: AY942107; AAX37275.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS00552; HTH_MERR_1; 1.
DR PROSITE; PS00937; HTH_MERR_2; 1.
KW DNA-binding.
SQ SEQUENCE 130 AA; 14496 MW; AB8EF07611855BDC CRC64;

Query Match 89.3%; Score 25; DB 2; Length 130;
Best Local Similarity 55.6%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
   ||||:
DB 103 VRDQNTQL 111

RESULT 36
Q4XMZ2_PLACH PRELIMINARY; PRT; 139 AA.
ID Q4XMZ2_PLACH
AC Q4XMZ2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PC000779.04.0;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=5825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAAJ01004543; CAH81720.1; -; Genomic_DNA.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 16249 MW; 5A31ABF4B835B1AE CRC64;

Query Match 89.3%; Score 25; DB 2; Length 139;
Best Local Similarity 55.6%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```

Qy      1 VRDQIXXXL 9
Db      96 IRDQIALL 104

RESULT 37
Q916T7_PSEAE PRELIMINARY;      PRT;      193 AA.
AC      Q916T7;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hypothetical protein.
GN      OrderedLocusNames=PA0201;
OS      Pseudomonas aeruginosa.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=287;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=ATCC 15692 / PA01;
RX      MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA      Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA      Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA      Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA      Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA      Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA      Reizer J., Sailer M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT      "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT      opportunistic pathogen.";
RL      Nature 406:959-964(2000).
DR      EMBL; AE004458; AAG03590.1; -; Genomic_DNA.
DR      PIR; D83620; D83620.
DR      InterPro; IPR010662; DUF1234.
DR      Pfam; PF06821; DUF1234; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 193 AA; 21734 MW; 1FDC933C4DD794BE CRC64;

Query Match      89.3%; Score 25; DB 2; Length 193;
Best Local Similarity 55.6%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VRDQIXXXL 9
Db      87 VRDQVRGAL 95

RESULT 38
Q65L06_BACLD PRELIMINARY;      PRT;      225 AA.
AC      Q65L06; Q62WF3;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Hypothetical DNA-binding protein, putative transcriptional
DE      regulator.
GN      OrderedLocusNames=BL03804, BL101354;
OS      Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=279010;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      PubMed=15383718; DOI=10.1159/000079829;
RX      Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA      Ehrenreich P., Baumeier S., Henne A., Liesegang H., Merkl R.,
RA      Ehrenreich A., Gottschalk G.;
RT      "The complete genome sequence of Bacillus licheniformis DSM13, an
RT      organism with great industrial potential.";
RL      J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;

```

```

RA      Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA      Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA      Olsen P.B., Raemussen M.D., Andersen J.T., Joergensen P.L.,
RA      Larsen T.S., Sorokin A., Bolotin A., Lepoint A., Galleron N.,
RA      Ehrlich S.D., Berka R.M.;
RT      "Complete genome sequence of the industrial bacterium Bacillus
RT      licheniformis and comparisons with closely related Bacillus species.";
RL      Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
DR      EMBL; AE017333; AAU40258.1; -; Genomic_DNA.
DR      EMBL; CP000002; AAU22905.1; -; Genomic_DNA.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      GO; GO:0006350; P:transcription; IEA.
DR      InterPro; IPR011711; GntR_C.
DR      InterPro; IPR000524; HTH_GntR.
DR      Pfam; PF07729; FCD; 1.
DR      Pfam; PF00392; GntR; 1.
DR      PRINTS; PR00035; HTHGNTR.
DR      SMART; SM00345; HTH_GNTR; 1.
DR      PROSITE; PS00949; HTH_GNTR; 1.
KW      Complete proteome; DNA-binding; Hypothetical protein; Transcription;
KW      Transcription regulation.
SQ      SEQUENCE 225 AA; 26425 MW; 7F60B5D7770477EB CRC64;

Query Match      89.3%; Score 25; DB 2; Length 225;
Best Local Similarity 66.7%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VRDQIXXXL 9
Db      11 VRDQIYQVL 19

RESULT 39
Q6VT06_9VIRU PRELIMINARY;      PRT;      229 AA.
ID      Q6VT06_9VIRU PRELIMINARY;
AC      Q6VT06;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Hypothetical protein orf11C.
GN      Name=orf11C;
OS      Vibrio parahaemolyticus phage VP16C.
OC      Viruses.
OX      NCBI_TaxID=238893;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      PubMed=14563879; DOI=10.1128/JB.185.21.6434-6447.2003;
RA      Seguritan V., Feng I.W., Rohwer F., Swift M., Segall A.M.;
RT      "Genome sequences of two closely related Vibrio parahaemolyticus
RT      phages, VP16T and VP16C.";
RL      J. Bacteriol. 185:6434-6447(2003).
DR      EMBL; AY328853; AAQ96542.1; -; Genomic_DNA.
KW      Hypothetical protein.
SQ      SEQUENCE 229 AA; 25407 MW; 38CFE17DC93CD5DA CRC64;

Query Match      89.3%; Score 25; DB 2; Length 229;
Best Local Similarity 55.6%; Pred. No. 6.8e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 VRDQIXXXL 9
Db      18 VRDQVAQIL 26

RESULT 40
P90912_CAEBL PRELIMINARY;      PRT;      243 AA.
ID      P90912_CAEBL PRELIMINARY;
AC      P90912;
DT      01-MAY-1997 (TrEMBLrel. 03, Created)
DT      01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```


Hypothetical protein K07A1.7.
 ORFNames=K07A1.7;
 Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology.";
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z81097; CAB03168.1; -; Genomic_DNA.
 DR PIR; T23381; T23381.
 DR Ensembl; K07A1.7; Caenorhabditis elegans.
 DR WormBase; WGenome00010614; K07A1.7.
 DR WormPep; K07A1.7; CE11852.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 243 AA; 27965 MW; 38042BB7692A2988 CRC64;
 Query Match 89.3%; Score 25; DB 2; Length 243;
 Best Local Similarity 55.6%; Pred. No. 7.2e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VRDQIXXXL 9
 DB 206 VRDQIXXXL 214
 RESULT 41
 EVGI_DROME STANDARD; PRT; 250 AA.
 AC Q9VSS7;
 ID 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0193 protein EVGI homolog.
 GN ORFNames=CG5280;
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Berkley;
 RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laake P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

```

OX NCBI_TaxID=7227;
RN
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungell C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celinker S.;
RL Submitted (MAY-2002) to the ENBL/GenBank/DBJ databases.
DR EMBL; AY113308; AAM29313.1; -; mRNA.
DR Ensembl; CG5280; Drosophila melanogaster.
DR FlyBase; FBgn0035952; CG5280.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR007914; UPR0193.
DR Pfam; PF05250; UPR0193; 1.
SQ SEQUENCE 250 AA; 28564 MW; 5EABD1FAF2E7225C CRC64;

Query Match 89.3%; Score 25; DB 2; Length 250;
Best Local Similarity 55.6%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VRDQIXXXL 9
Db 220 IRDQIAERL 228
|||||
1 220 IRDQIAERL 228

RESULT 43
ID Q72MR9 LEPIC PRELIMINARY; PRT; 291 AA.
AC Q72MR9
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=L1C1323;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroseae; Leptospira.
OX NCBI_TaxID=44275;
RN
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Piocruz L1-130;
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorcy H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
RA Kuramae E.S., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,
RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomes of two Leptospira interrogans serovars reveals novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AE017299; AAS71669.1; -; Genomic DNA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR Complete proteome.
KW PRINTS; PR01590; HTHFIS.
SQ SEQUENCE 291 AA; 33139 MW; 28CA78664F554E1D CRC64;

Query Match 89.3%; Score 25; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 VRDQIXXXL 9
Db 266 VRDQIRKDL 274
|||||
1 266 VRDQIRKDL 274

```

```

RESULT 44
Q8EZE2 LEPIN PRELIMINARY; PRT; 291 AA.
AC Q8EZE2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=LA3912;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroseae; Leptospira.
OX NCBI_TaxID=173;
RN
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AE011546; AAM51110.1; -; Genomic DNA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011989; ARM-like.
DR InterPro; IPR002197; HTH_Fis.
DR PRINTS; PR01590; HTHFIS.
KW Complete proteome.
SQ SEQUENCE 291 AA; 33125 MW; E118E8639BBDE15 CRC64;

Query Match 89.3%; Score 25; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 VRDQIXXXL 9
Db 266 VRDQIRKDL 274
|||||
1 266 VRDQIRKDL 274

RESULT 45
Q8G822 BIFLO PRELIMINARY; PRT; 296 AA.
AC Q8G822
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Possible TetR-type transcriptional regulator.
GN OrderedLocustNames=BL0061;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schell M.A., Karimantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014295; AAN23927.1; -; Genomic DNA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

```

```

DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PROSITE; PS00977; HTH_TETR_2; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33740 MW; 4522D8B92DE048D6 CRC64;

Query Match      89.3%; Score 25; DB 2; Length 296;
Best Local Similarity 66.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 199 VRDQIEDDL 207

RESULT 46
PERN ARMURU
ID PERN ARMURU STANDARD; PRT; 327 AA.
AC Q42517;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Peroxidase N precursor (EC 1.11.1.7) (Neutral peroxidase).
GN Name=HRPN;
OS Armoracia rusticana (Horseradish) (Armoracia laphatifolia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Armoracia.
OX NCBI_TaxID=3704;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Root;
RX MEDLINE=91159476; PubMed=2001399;
RA Bartonek-Roxa E., Eriksson H., Mattiasson B.;
RT "The cDNA sequence of a neutral horseradish peroxidase.";
RL Biochim. Biophys. Acta 1088:245-250(1991).
CC -!- FUNCTION: Removal of H(2)O(2), oxidation of toxic reductants,
CC biosynthesis and degradation of lignin, suberization, auxin
CC catabolism, response to environmental stresses such as wounding,
CC pathogen attack and oxidative stress. These functions might be
CC dependent on each isozyme/isoform in each plant tissue.
CC -!- CATALYTIC ACTIVITY: Donor + H(2)O(2) = oxidized donor + 2 H(2)O.
CC -!- COFACTOR: Binds 2 calcium ions per subunit.
CC -!- COFACTOR: Binds 1 heme B (iron-protophyrin IX) group per
CC subunit.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the peroxidase family. Classical plant
CC (class III) peroxidase subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X57564; CAA04796.1; -; mRNA.
CC FIR; S14268; S14268.
CC DR HSP; Q39034; 1QGJ.
CC DR SNR; Q42517; 29-327.
CC DR InterPro; IPR000823; Plant_peroxidase.
CC DR InterPro; IPR002016; Peroxidase.
CC DR Pfam; PF00141; peroxidase; 1.
CC DR PRINTS; PR00458; PEROXIDASE.
CC DR PROSITE; PS00461; PEROXIDASE.
CC DR PROSITE; PS00435; PEROXIDASE_1; 1.
CC DR PROSITE; PS00436; PEROXIDASE_2; 1.
CC DR PROSITE; PS00873; PEROXIDASE_4; 1.
CC DR Calcium; Glycoprotein; Heme; Hydrogen peroxide; Iron; Metal-binding;
KW Multigene family; Oxidoreductase; Peroxidase;
KW Pyrrolidone carboxylic acid; Signal.
FT SIGNAL 1 28 Potential.
FT CHAIN 29 327 Peroxidase N.

```

```

FT ACT_SITE 70 70 Proton acceptor (By similarity).
FT METAL 71 71 Calcium 1 (By similarity).
FT METAL 74 74 Calcium 1 (via carbonyl oxygen) (By
FT METAL 76 76 Calcium 1 (via carbonyl oxygen) (By
FT METAL 78 78 Calcium 1 (By similarity).
FT METAL 80 80 Calcium 1 (By similarity).
FT METAL 193 193 Iron (heme axial ligand) (By similarity).
FT METAL 194 194 Calcium 2 (By similarity).
FT METAL 245 245 Calcium 2 (By similarity).
FT METAL 248 248 Calcium 2 (By similarity).
FT METAL 253 253 Calcium 2 (By similarity).
FT BINDING 163 163 Substrate (By similarity).
FT SITE 66 66 Transition state stabilizer (By
FT MOD_RES 29 29 Pyrrolidone carboxylic acid (By
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 182 182 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 209 209 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 239 239 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 247 247 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 281 281 N-linked (GlcNAc...) (Potential).
FT DISULFID 39 116 By similarity.
FT DISULFID 72 77 By similarity.
FT DISULFID 122 323 By similarity.
FT DISULFID 200 232 By similarity.
SQ SEQUENCE 327 AA; 35126 MW; 5C427EBDD0A2CCDF CRC64;

Query Match      89.3%; Score 25; DB 1; Length 327;
Best Local Similarity 55.6%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 46 VRDQVKIAL 54

RESULT 47
Q61Y40 CAEBR
ID Q61Y40 CAEBR PRELIMINARY; PRT; 358 AA.
AC Q61Y40;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG03720.
GN Name=CBG03720;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
-!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAC01000016; CAB60166.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 358 AA; 40922 MW; 07DFE0D15316381C CRC64;

Query Match      89.3%; Score 25; DB 2; Length 358;
Best Local Similarity 55.6%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VRDQIXXXL 9
DB 309 VRDQLATQL 317

RESULT 48

```

```

DR EMBL; AE016806; AAO11189.1; -; Genomic_DNA.
DR GO; GO:0042597; C:periplasmic space; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR DR InterPro; IPR010869; DUF1501.
DR DR InterPro; IPR006311; Tat.
DR PFam; PF07394; DUF1501.1.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
DR Complete proteome.
SQ SEQUENCE 437 AA; 47709 MW; 6250B6240B304753 CRC64;

Query Match      89.3%; Score 25; DB 2; Length 437;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 243 IRDQLSAAL 251

RESULT 50
Q7MLK9 VIBVY
ID Q7MLK9 VIBVY PRELIMINARY; PRT; 437 AA.
AC Q7MLK9;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein VV1418.
GN OrderedLocusNames=VV1418;
OS Vibrio vulnificus (strain XJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14658965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC94182.1; -; Genomic_DNA.
DR InterPro; IPR010869; DUF1501.
DR InterPro; IPR006311; Tat.
DR PFam; PF07394; DUF1501.1.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 437 AA; 47755 MW; 117B4F65E97D43E8 CRC64;

Query Match      89.3%; Score 25; DB 2; Length 437;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VRDQIXXXL 9
Db 243 IRDQLSAAL 251

Search completed: May 12, 2006, 10:52:58
Job time : 90.6923 secs

```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 43.5897 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-34

Perfect score: 34

Sequence: 1 TLREWL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21.*

1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	6	3 AAB16978	Aab16978 TPO-mimet
2	34	100.0	6	5 ABB72864	Abb72864 TPO mimet
3	34	100.0	6	6 ABG71746	Abg71746 TPO recep
4	34	100.0	6	7 ADJ73015	Adj73015 TPO mimet
5	34	100.0	6	8 ADJ52650	Adj52650 CH1 delet
6	34	100.0	6	8 ADJ51611	Adj51611 CH1 delet
7	34	100.0	9	7 ADJ73040	Adj73040 TPO mimet
8	34	100.0	9	7 ADJ73037	Adj73037 TPO mimet
9	34	100.0	9	7 ADJ73041	Adj73041 TPO mimet
10	34	100.0	9	7 ADJ73039	Adj73039 TPO mimet
11	34	100.0	9	7 ADJ73038	Adj73038 TPO mimet
12	34	100.0	10	2 AAU09471	Aau09471 Thrombopo
13	34	100.0	10	2 AAU09566	Aau09566 Thrombopo
14	34	100.0	10	2 AAU09622	Aau09622 Thrombopo
15	34	100.0	10	2 AAU09671	Aau09671 Thrombopo
16	34	100.0	10	3 AAB17000	Aab17000 TPO-mimet
17	34	100.0	10	4 AAU25841	Aau25841 Human thr
18	34	100.0	10	4 AAU25936	Aau25936 Human thr
19	34	100.0	10	5 ABB72886	Abb72886 TPO mimet
20	34	100.0	10	8 ADJ52672	Adj52672 CH1 delet
21	34	100.0	10	8 ADJ51633	Adj51633 CH1 delet
22	34	100.0	11	3 AAB17001	Aab17001 TPO-mimet
23	34	100.0	11	5 ABB72887	Abb72887 TPO mimet
24	34	100.0	11	8 ADJ52673	Adj52673 CH1 delet

25	34	100.0	11	8 ADJ51634	Adj51634 CH1 delet
26	34	100.0	12	2 AAU09476	Aau09476 Thrombopo
27	34	100.0	12	2 AAU09581	Aau09581 Thrombopo
28	34	100.0	12	2 AAU09577	Aau09577 Thrombopo
29	34	100.0	12	2 AAU09524	Aau09524 Thrombopo
30	34	100.0	12	2 AAU09524	Aau09524 Thrombopo
31	34	100.0	12	2 AAU09524	Aau09524 Thrombopo
32	34	100.0	12	2 AAU09524	Aau09524 Thrombopo
33	34	100.0	12	2 AAU09524	Aau09524 Thrombopo
34	34	100.0	12	3 AAB17002	Aab17002 TPO-mimet
35	34	100.0	12	3 AAB16993	Aab16993 TPO-mimet
36	34	100.0	12	4 AAU25846	Aau25846 Human thr
37	34	100.0	12	4 AAU25951	Aau25951 Human thr
38	34	100.0	12	4 AAU25894	Aau25894 Human thr
39	34	100.0	12	4 AAU25947	Aau25947 Human thr
40	34	100.0	12	5 ABB72888	Abb72888 TPO mimet
41	34	100.0	12	5 ABB72879	Abb72879 TPO mimet
42	34	100.0	12	7 ADJ73030	Adj73030 TPO mimet
43	34	100.0	12	8 ADJ52665	Adj52665 CH1 delet
44	34	100.0	12	8 ADJ52674	Adj52674 CH1 delet
45	34	100.0	12	8 ADJ51626	Adj51626 CH1 delet
46	34	100.0	12	8 ADJ51635	Adj51635 CH1 delet
47	34	100.0	13	2 AAU36732	Aau36732 Thrombopo
48	34	100.0	13	3 AAB17015	Aab17015 TPO-mimet
49	34	100.0	13	3 AAB17003	Aab17003 TPO-mimet
50	34	100.0	13	5 ABB72889	Abb72889 TPO mimet
51	34	100.0	13	7 ADJ73054	Adj73054 TPO mimet
52	34	100.0	13	7 ADJ73052	Adj73052 TPO mimet
53	34	100.0	13	7 ADJ73056	Adj73056 TPO mimet
54	34	100.0	13	7 ADJ73053	Adj73053 TPO mimet
55	34	100.0	13	7 ADJ73055	Adj73055 TPO mimet
56	34	100.0	13	8 ADJ52687	Adj52687 CH1 delet
57	34	100.0	13	8 ADJ52675	Adj52675 CH1 delet
58	34	100.0	13	8 ADJ51636	Adj51636 CH1 delet
59	34	100.0	13	8 ADJ51648	Adj51648 CH1 delet
60	34	100.0	14	3 AAB17017	Aab17017 TPO-mimet
61	34	100.0	14	3 AAB17016	Aab17016 TPO-mimet
62	34	100.0	14	3 AAB16969	Aab16969 TPO-mimet
63	34	100.0	14	3 AAB17004	Aab17004 TPO-mimet
64	34	100.0	14	5 ABB72903	Abb72903 TPO mimet
65	34	100.0	14	5 ABB72855	Abb72855 TPO mimet
66	34	100.0	14	5 ABB72890	Abb72890 TPO mimet
67	34	100.0	14	5 ABB72902	Abb72902 TPO mimet
68	34	100.0	14	7 ADJ73005	Adj73005 TPO mimet
69	34	100.0	14	8 ADJ52676	Adj52676 CH1 delet
70	34	100.0	14	8 ADJ52689	Adj52689 CH1 delet
71	34	100.0	14	8 ADJ52640	Adj52640 CH1 delet
72	34	100.0	14	8 ADJ52688	Adj52688 CH1 delet
73	34	100.0	14	8 ADJ51601	Adj51601 CH1 delet
74	34	100.0	14	8 ADJ51650	Adj51650 CH1 delet
75	34	100.0	14	8 ADJ51649	Adj51649 CH1 delet
76	34	100.0	14	8 ADJ51637	Adj51637 CH1 delet
77	34	100.0	15	3 AAB17018	Aab17018 TPO-mimet
78	34	100.0	15	5 ABB72904	Abb72904 TPO mimet
79	34	100.0	15	8 ADJ52691	Adj52691 CH1 delet
80	34	100.0	15	8 ADJ52690	Adj52690 CH1 delet
81	34	100.0	15	8 ADJ51652	Adj51652 CH1 delet
82	34	100.0	15	8 ADJ51651	Adj51651 CH1 delet
83	34	100.0	16	2 AAU09464	Aau09464 Thrombopo
84	34	100.0	16	3 AAB17019	Aab17019 TPO-mimet
85	34	100.0	16	3 AAB17019	Aab17019 TPO-mimet
86	34	100.0	16	5 ABB72905	Abb72905 TPO mimet
87	34	100.0	16	7 ADJ73057	Adj73057 TPO mimet
88	34	100.0	16	8 ADJ52692	Adj52692 CH1 delet
89	34	100.0	16	8 ADJ51653	Adj51653 CH1 delet
90	34	100.0	18	2 AAU09460	Aau09460 Thrombopo
91	34	100.0	18	2 AAU09583	Aau09583 Thrombopo
92	34	100.0	18	2 AAU09498	Aau09498 Thrombopo
93	34	100.0	18	2 AAU09486	Aau09486 Thrombopo
94	34	100.0	18	2 AAU09486	Aau09486 Thrombopo
95	34	100.0	18	2 AAU09486	Aau09486 Thrombopo
96	34	100.0	18	2 AAU09486	Aau09486 Thrombopo
97	34	100.0	18	2 AAU09486	Aau09486 Thrombopo

98	34	100.0	18	2	AAW33027	Aaw33027	Thrombopo	171	34	100.0	422	3	AAW96815	Aaw96815	A. thalia
99	34	100.0	18	2	AAW36652	Aaw36652	Thrombopo	172	34	100.0	455	3	AAW59392	Aaw59392	Arabidops
100	34	100.0	18	3	AAW17026	Aad17026	TPO-mimet	173	34	100.0	486	3	AAW59391	Aaw59391	Arabidops
101	34	100.0	18	4	AAU25856	Aau25856	Human thr	174	34	100.0	496	9	ABM91663	Abm91663	M. xanthu
102	34	100.0	18	4	AAU25835	Aau25835	Human thr	175	34	100.0	530	3	AAW59390	Aaw59390	Arabidops
103	34	100.0	18	4	AAU25868	Aau25868	Human thr	176	34	94.1	9	2	AAW35429	Aaw35429	Thrombopo
104	34	100.0	18	4	AAU25953	Aau25953	Human thr	177	34	94.1	9	4	AAU26003	Aau26003	Human thr
105	34	100.0	18	4	AAU25837	Aau25837	Human thr	178	34	94.1	10	2	AAW35427	Aaw35427	Thrombopo
106	34	100.0	18	4	AAU25824	Aau25824	Human thr	179	34	94.1	10	4	AAU26002	Aau26002	Human thr
107	34	100.0	18	4	AAU25871	Aau25871	Human thr	180	34	94.1	11	2	AAW35425	Aaw35425	Thrombopo
108	34	100.0	18	5	ABB72912	Abb72912	TPO mimet	181	34	94.1	11	4	AAU26001	Aau26001	Human thr
109	34	100.0	18	7	ADJ73064	Adj73064	TPO mimet	182	34	94.1	12	2	AAW35423	Aaw35423	Thrombopo
110	34	100.0	18	7	ADN59664	Adn59664	Thrombopo	183	34	94.1	12	4	AAU26000	Aau26000	Human thr
111	34	100.0	18	7	ADN59672	Adn59672	Thrombopo	184	34	94.1	12	8	ADM72530	Adm72530	TPO mimet
112	34	100.0	18	7	ADN59658	Adn59658	Thrombopo	185	34	94.1	13	2	AAW09467	Aaw09467	Thrombopo
113	34	100.0	18	7	ADN59675	Adn59675	Thrombopo	186	34	94.1	13	2	AAW35399	Aaw35399	Thrombopo
114	34	100.0	18	7	ADN59670	Adn59670	Thrombopo	187	34	94.1	13	2	AAW35404	Aaw35404	Thrombopo
115	34	100.0	18	7	ADN59668	Adn59668	Thrombopo	188	34	94.1	13	2	AAW35417	Aaw35417	Thrombopo
116	34	100.0	18	7	ADN59669	Adn59669	Thrombopo	189	34	94.1	13	2	AAW33033	Aaw33033	Thrombopo
117	34	100.0	18	7	ADN59657	Adn59657	Thrombopo	190	34	94.1	13	2	AAW35413	Aaw35413	Thrombopo
118	34	100.0	18	7	ADN59673	Adn59673	Thrombopo	191	34	94.1	13	2	AAW35406	Aaw35406	Thrombopo
119	34	100.0	18	7	ADN59667	Adn59667	Thrombopo	192	34	94.1	13	2	AAW35405	Aaw35405	Thrombopo
120	34	100.0	18	8	ADJ52699	Adj52699	CHI delet	193	34	94.1	13	2	AAW35422	Aaw35422	Thrombopo
121	34	100.0	18	8	ADJ51660	Adj51660	CHI delet	194	34	94.1	13	2	AAW35397	Aaw35397	Thrombopo
122	34	100.0	19	4	AAU25836	Aau25836	Human thr	195	34	94.1	13	4	AAU26039	Aau26039	Human thr
123	34	100.0	22	7	ADN59840	Adn59840	TMP pepti	196	34	94.1	13	4	AAU26041	Aau26041	Human thr
124	34	100.0	22	7	ADN59824	Adn59824	TMP pepti	197	34	94.1	13	4	AAU25994	Aau25994	Human thr
125	34	100.0	22	7	ADN59836	Adn59836	TMP pepti	198	34	94.1	13	4	AAU25997	Aau25997	Human thr
126	34	100.0	22	7	ADN59835	Adn59835	TMP pepti	199	34	94.1	13	4	AAU25991	Aau25991	Human thr
127	34	100.0	22	7	ADN59831	Adn59831	TMP pepti	200	34	94.1	13	4	AAU25990	Aau25990	Human thr
128	34	100.0	22	7	ADN59825	Adn59825	TMP pepti	201	34	94.1	13	4	AAU25984	Aau25984	Human thr
129	34	100.0	22	7	ADN59839	Adn59839	TMP pepti	202	34	94.1	13	8	ADM72529	Adm72529	TPO mimet
130	34	100.0	22	7	ADN59834	Adn59834	TMP pepti	203	34	94.1	13	8	ADM72528	Adm72528	TPO mimet
131	34	100.0	23	7	ADN59775	Adn59775	Peptide-v	204	34	94.1	14	2	AAW09466	Aaw09466	Thrombopo
132	34	100.0	23	7	ADN59778	Adn59778	Peptide-v	205	34	94.1	14	2	AAW09462	Aaw09462	Thrombopo
133	34	100.0	23	7	ADN59793	Adn59793	Peptide-v	206	34	94.1	14	2	AAW09465	Aaw09465	Thrombopo
134	34	100.0	23	7	ADN59796	Adn59796	Peptide-v	207	34	94.1	14	2	AAW09482	Aaw09482	Thrombopo
135	34	100.0	25	7	ADN59726	Adn59726	Thrombopo	208	34	94.1	14	2	AAW33031	Aaw33031	Thrombopo
136	34	100.0	25	7	ADN59696	Adn59696	Thrombopo	209	34	94.1	14	2	AAW35412	Aaw35412	Thrombopo
137	34	100.0	25	7	ADN59730	Adn59730	Thrombopo	210	34	94.1	14	2	AAW35407	Aaw35407	Thrombopo
138	34	100.0	25	7	ADN59710	Adn59710	Thrombopo	211	34	94.1	14	2	AAW35408	Aaw35408	Thrombopo
139	34	100.0	25	7	ADN59698	Adn59698	Thrombopo	212	34	94.1	14	2	AAW35398	Aaw35398	Thrombopo
140	34	100.0	25	7	ADN59716	Adn59716	Thrombopo	213	34	94.1	14	2	AAW36633	Aaw36633	Thrombopo
141	34	100.0	25	7	ADN59718	Adn59718	Thrombopo	214	34	94.1	14	2	AAW33029	Aaw33029	Thrombopo
142	34	100.0	25	7	ADN59740	Adn59740	Thrombopo	215	34	94.1	14	2	AAW35396	Aaw35396	Thrombopo
143	34	100.0	25	7	ADN59720	Adn59720	Thrombopo	216	34	94.1	14	2	AAW35401	Aaw35401	Thrombopo
144	34	100.0	25	7	ADN59744	Adn59744	Thrombopo	217	34	94.1	14	2	AAW35403	Aaw35403	Thrombopo
145	34	100.0	36	7	ADN59763	Adn59763	Peptide-v	218	34	94.1	14	2	AAW36647	Aaw36647	Thrombopo
146	34	100.0	36	7	ADN59766	Adn59766	Peptide-v	219	34	94.1	14	2	AAW35400	Aaw35400	Thrombopo
147	34	100.0	36	7	ADN59756	Adn59756	Peptide-v	220	34	94.1	14	2	AAW35402	Aaw35402	Thrombopo
148	34	100.0	41	7	ADN59769	Adn59769	Peptide-v	221	34	94.1	14	2	AAW33032	Aaw33032	Thrombopo
149	34	100.0	41	7	ADN59816	Adn59816	Peptide-v	222	34	94.1	14	2	AAW66732	Aaw66732	Peptide c
150	34	100.0	41	7	ADN59772	Adn59772	Peptide-v	223	34	94.1	14	3	AAW17014	Aab17014	TPO-mimet
151	34	100.0	42	7	ADN59818	Adn59818	Peptide-v	224	34	94.1	14	4	AAU25987	Aau25987	Human thr
152	34	100.0	46	7	ADN59781	Adn59781	Peptide-v	225	34	94.1	14	4	AAU25826	Aau25826	Human thr
153	34	100.0	46	7	ADN59787	Adn59787	Peptide-v	226	34	94.1	14	4	AAU25993	Aau25993	Human thr
154	34	100.0	46	7	ADN59790	Adn59790	Peptide-v	227	34	94.1	14	4	AAU25852	Aau25852	Human thr
155	34	100.0	46	7	ADN59784	Adn59784	Peptide-v	228	34	94.1	14	4	AAU25989	Aau25989	Human thr
156	34	100.0	160	4	ABB58297	Abb58297	Drosophil	229	34	94.1	14	4	AAU25983	Aau25983	Human thr
157	34	100.0	226	3	AAW58613	Aaw58613	Protein r	230	34	94.1	14	4	AAU25985	Aau25985	Human thr
158	34	100.0	235	3	AAW57785	Aaw57785	Arabidops	231	34	94.1	14	4	AAU25995	Aau25995	Human thr
159	34	100.0	235	3	AAW57786	Aaw57786	Arabidops	232	34	94.1	14	4	AAU25992	Aau25992	Human thr
160	34	100.0	235	3	AAW55516	Aaw55516	Arabidops	233	34	94.1	14	4	AAU25986	Aau25986	Human thr
161	34	100.0	235	3	AAW57791	Aaw57791	Arabidops	234	34	94.1	14	4	AAU26040	Aau26040	Human thr
162	34	100.0	243	7	ADJ69556	Adj69556	Human hea	235	34	94.1	14	4	AAU25866	Aau25866	Human thr
163	34	100.0	302	3	AAW55515	Aaw55515	Arabidops	236	34	94.1	14	4	AAU25988	Aau25988	Human thr
164	34	100.0	304	7	ADP07784	Adp07784	Bacterial	237	34	94.1	14	5	ABB72900	Abb72900	TPO mimet
165	34	100.0	313	6	ABU89792	Abu89792	Novel hum	238	34	94.1	14	7	ADJ73051	Adj73051	TPO mimet
166	34	100.0	313	8	ADO09332	Ado09332	Novel hum	239	34	94.1	14	8	ADJ52686	Adj52686	CHI delet
167	34	100.0	325	7	ADB75420	Adb75420	Prostate	240	34	94.1	14	8	ADJ51647	Adj51647	CHI delet
168	34	100.0	326	5	ABP43878	Abp43878	LRP16 pro	241	34	94.1	18	2	AAW09596	Aaw09596	Thrombopo
169	34	100.0	408	3	AAW57790	Aaw57790	Arabidops	242	34	94.1	18	2	AAW09456	Aaw09456	Thrombopo
170	34	100.0	413	3	AAW57789	Aaw57789	Arabidops	243	34	94.1	18	2	AAW33023	Aaw33023	Thrombopo

244	32	94.1	18	2	AAW36747	Thrombopo	317	31	91.2	13	2	AAW36779	Thrombopo
245	32	94.1	18	2	AAW36748	Thrombopo	318	31	91.2	13	2	AAW36783	Thrombopo
246	32	94.1	18	3	AAW36748	TPO-mimet	319	31	91.2	13	3	AAW36783	TPO-mimet
247	32	94.1	18	3	AAW36748	TPO-mimet	320	31	91.2	13	3	AAW36783	TPO-mimet
248	32	94.1	18	4	AAW36748	Human thr	321	31	91.2	13	4	AAW36783	Human thr
249	32	94.1	18	4	AAW36748	Human thr	322	31	91.2	13	4	AAW36783	Human thr
250	32	94.1	18	5	AAW36748	Human thr	323	31	91.2	13	5	AAW36783	Human thr
251	32	94.1	18	5	AAW36748	TPO mimet	324	31	91.2	13	5	AAW36783	TPO mimet
252	32	94.1	18	8	AAW36748	CH1 delet	325	31	91.2	13	8	AAW36783	CH1 delet
253	32	94.1	18	8	AAW36748	CH1 delet	326	31	91.2	13	8	AAW36783	CH1 delet
254	32	94.1	18	8	AAW36748	Thrombopo	327	31	91.2	13	8	AAW36783	Thrombopo
255	32	94.1	19	2	AAW36748	Thrombopo	328	31	91.2	13	2	AAW36783	Thrombopo
256	32	94.1	19	4	AAW36748	Human thr	329	31	91.2	13	4	AAW36783	Human thr
257	32	94.1	25	4	AAW36748	Human thr	330	31	91.2	13	4	AAW36783	Human thr
258	32	94.1	25	8	AAW36748	TPO mimet	331	31	91.2	13	8	AAW36783	TPO mimet
259	32	94.1	109	3	AAW36748	Lung canc	332	31	91.2	13	3	AAW36783	Lung canc
260	32	94.1	169	7	AAW36748	Rice abio	333	31	91.2	13	7	AAW36783	Rice abio
261	32	94.1	176	5	AAW36748	Rice Fr h	334	31	91.2	13	5	AAW36783	Rice Fr h
262	32	94.1	176	8	AAW36748	Rice Fr h	335	31	91.2	13	8	AAW36783	Rice Fr h
263	32	94.1	313	3	AAW36748	Arabidops	336	31	91.2	13	3	AAW36783	Arabidops
264	32	94.1	313	3	AAW36748	Arabidops	337	31	91.2	13	3	AAW36783	Arabidops
265	32	94.1	313	8	AAW36748	Plant ful	338	31	91.2	13	8	AAW36783	Plant ful
266	32	94.1	413	8	AAW36748	Plant ful	339	31	91.2	13	8	AAW36783	Plant ful
267	32	94.1	413	8	AAW36748	Plant ful	340	31	91.2	13	8	AAW36783	Plant ful
268	32	94.1	450	3	AAW36748	Arabidops	341	31	91.2	13	3	AAW36783	Arabidops
269	32	94.1	499	3	AAW36748	Arabidops	342	31	91.2	13	3	AAW36783	Arabidops
270	32	94.1	537	3	AAW36748	Arabidops	343	31	91.2	13	3	AAW36783	Arabidops
271	32	94.1	569	3	AAW36748	Arabidops	344	31	91.2	13	3	AAW36783	Arabidops
272	32	94.1	570	8	AAW36748	Arabidops	345	31	91.2	13	8	AAW36783	Arabidops
273	32	94.1	577	3	AAW36748	Protein e	346	31	91.2	13	3	AAW36783	Protein e
274	32	94.1	834	6	AAW36748	Microbial	347	31	91.2	13	6	AAW36783	Microbial
275	32	94.1	834	6	AAW36748	Microbial	348	31	91.2	13	6	AAW36783	Microbial
276	31	91.2	8	2	AAW36748	Thrombopo	349	31	91.2	14	2	AAW36783	Thrombopo
277	31	91.2	8	5	AAW36748	Human thr	350	31	91.2	14	5	AAW36783	Human thr
278	31	91.2	8	5	AAW36748	TPO mimet	351	31	91.2	14	5	AAW36783	TPO mimet
279	31	91.2	8	8	AAW36748	TPO mimet	352	31	91.2	14	8	AAW36783	TPO mimet
280	31	91.2	8	9	AAW36748	Agonist T	353	31	91.2	14	9	AAW36783	Agonist T
281	31	91.2	8	9	AAW36748	Agonist T	354	31	91.2	14	9	AAW36783	Agonist T
282	31	91.2	10	2	AAW36748	Thrombopo	355	31	91.2	14	2	AAW36783	Thrombopo
283	31	91.2	10	2	AAW36748	Thrombopo	356	31	91.2	14	2	AAW36783	Thrombopo
284	31	91.2	10	2	AAW36748	Thrombopo	357	31	91.2	14	2	AAW36783	Thrombopo
285	31	91.2	10	2	AAW36748	Thrombopo	358	31	91.2	14	2	AAW36783	Thrombopo
286	31	91.2	10	3	AAW36748	TPO-mimet	359	31	91.2	14	3	AAW36783	TPO-mimet
287	31	91.2	10	3	AAW36748	TPO-mimet	360	31	91.2	14	3	AAW36783	TPO-mimet
288	31	91.2	10	4	AAW36748	Human thr	361	31	91.2	14	4	AAW36783	Human thr
289	31	91.2	10	4	AAW36748	Human thr	362	31	91.2	14	4	AAW36783	Human thr
290	31	91.2	10	5	AAW36748	TPO mimet	363	31	91.2	14	5	AAW36783	TPO mimet
291	31	91.2	10	5	AAW36748	TPO mimet	364	31	91.2	14	5	AAW36783	TPO mimet
292	31	91.2	10	7	AAW36748	TPO mimet	365	31	91.2	14	7	AAW36783	TPO mimet
293	31	91.2	10	7	AAW36748	TPO mimet	366	31	91.2	14	7	AAW36783	TPO mimet
294	31	91.2	10	8	AAW36748	CH1 delet	367	31	91.2	14	8	AAW36783	CH1 delet
295	31	91.2	10	8	AAW36748	CH1 delet	368	31	91.2	14	8	AAW36783	CH1 delet
296	31	91.2	10	8	AAW36748	CH1 delet	369	31	91.2	14	8	AAW36783	CH1 delet
297	31	91.2	10	8	AAW36748	CH1 delet	370	31	91.2	14	8	AAW36783	CH1 delet
298	31	91.2	11	2	AAW36791	Thrombopo	371	31	91.2	14	2	AAW36791	Thrombopo
299	31	91.2	12	2	AAW36791	Thrombopo	372	31	91.2	14	2	AAW36791	Thrombopo
300	31	91.2	12	2	AAW36791	Thrombopo	373	31	91.2	14	2	AAW36791	Thrombopo
301	31	91.2	12	2	AAW36791	Thrombopo	374	31	91.2	14	2	AAW36791	Thrombopo
302	31	91.2	12	2	AAW36791	Thrombopo	375	31	91.2	14	2	AAW36791	Thrombopo
303	31	91.2	12	2	AAW36791	Thrombopo	376	31	91.2	14	2	AAW36791	Thrombopo
304	31	91.2	12	2	AAW36791	Thrombopo	377	31	91.2	14	2	AAW36791	Thrombopo
305	31	91.2	12	2	AAW36791	Thrombopo	378	31	91.2	14	2	AAW36791	Thrombopo
306	31	91.2	12	2	AAW36791	Thrombopo	379	31	91.2	14	2	AAW36791	Thrombopo
307	31	91.2	12	3	AAW36791	Human thr	380	31	91.2	14	3	AAW36791	Human thr
308	31	91.2	12	4	AAW36791	Human thr	381	31	91.2	14	4	AAW36791	Human thr
309	31	91.2	12	4	AAW36791	Human thr	382	31	91.2	14	4	AAW36791	Human thr
310	31	91.2	12	4	AAW36791	Human thr	383	31	91.2	14	4	AAW36791	Human thr
311	31	91.2	12	4	AAW36791	Human thr	384	31	91.2	14	4	AAW36791	Human thr
312	31	91.2	12	5	AAW36791	TPO mimet	385	31	91.2	14	5	AAW36791	TPO mimet
313	31	91.2	12	7	AAW36791	TPO mimet	386	31	91.2	14	7	AAW36791	TPO mimet
314	31	91.2	12	8	AAW36791	CH1 delet	387	31	91.2	14	8	AAW36791	CH1 delet
315	31	91.2	12	8	AAW36791	CH1 delet	388	31	91.2	14	8	AAW36791	CH1 delet
316	31	91.2	13	2	AAW36792	Thrombopo	389	31	91.2	14	2	AAW36792	Thrombopo

390	31	91.2	14	9	ADU70206	Adu70206	Thrombopo	463	31	91.2	18	3	AAB16957	Aab16957	PEGylated
391	31	91.2	14	9	ADU75982	Adu75982	Peptide-b	464	31	91.2	18	3	AAB16956	Aab16956	PEGylated
392	31	91.2	14	9	ADU75979	Adu75979	Thrombopo	465	31	91.2	18	3	AAB17024	Aab17024	TPO-mimet
393	31	91.2	14	9	ADV44319	Adv44319	Agonist T	466	31	91.2	18	3	AAB17025	Aab17025	TPO-mimet
394	31	91.2	14	9	ABB12792	Abb12792	TPO mimet	467	31	91.2	18	4	AAU25872	Aau25872	Human thr
395	31	91.2	15	2	AAW35416	Aaw35416	Thrombopo	468	31	91.2	18	4	AAU25867	Aau25867	Human thr
396	31	91.2	15	2	AAW36784	Aaw36784	Thrombopo	469	31	91.2	18	4	AAU25869	Aau25869	Human thr
397	31	91.2	15	2	AAW36780	Aaw36780	Thrombopo	470	31	91.2	18	4	AAU25823	Aau25823	Human thr
398	31	91.2	15	2	AAW36776	Aaw36776	Thrombopo	471	31	91.2	18	4	AAU25965	Aau25965	Human thr
399	31	91.2	15	2	AAW66731	Aaw66731	Peptide c	472	31	91.2	18	4	AAU25962	Aau25962	Human thr
400	31	91.2	15	2	AAW66717	Aaw66717	Peptide c	473	31	91.2	18	5	ABB72910	Abb72910	TPO mimet
401	31	91.2	15	2	AAW66718	Aaw66718	Peptide c	474	31	91.2	18	5	ABB72911	Abb72911	TPO mimet
402	31	91.2	15	2	AAW66714	Aaw66714	Peptide c	475	31	91.2	18	5	ABP51687	Abp51687	TPO mimet
403	31	91.2	15	2	AAW66721	Aaw66721	Peptide c	476	31	91.2	18	5	ABP51689	Abp51689	TPO mimet
404	31	91.2	15	2	AAW66712	Aaw66712	Peptide c	477	31	91.2	18	5	ABP51688	Abp51688	TPO mimet
405	31	91.2	15	3	AAAB20684	Aaab20684	Thrombocy	478	31	91.2	18	5	ABP51677	Abp51677	TPO mimet
406	31	91.2	15	4	AAU25992	Aau25992	Human thr	479	31	91.2	18	5	ABP51686	Abp51686	TPO mimet
407	31	91.2	15	4	AAU25996	Aau25996	Human thr	480	31	91.2	18	5	ABP51674	Abp51674	TPO mimet
408	31	91.2	15	4	AAU26026	Aau26026	Human thr	481	31	91.2	18	5	ABP51693	Abp51693	TPO mimet
409	31	91.2	15	4	AAU26011	Aau26011	Human thr	482	31	91.2	18	5	ABP51684	Abp51684	TPO mimet
410	31	91.2	15	4	AAU26020	Aau26020	Human thr	483	31	91.2	18	5	ABP51683	Abp51683	TPO mimet
411	31	91.2	15	4	AAU25831	Aau25831	Human thr	484	31	91.2	18	5	ABP51685	Abp51685	TPO mimet
412	31	91.2	15	4	AAU26023	Aau26023	Human thr	485	31	91.2	18	5	ABP51691	Abp51691	TPO mimet
413	31	91.2	15	4	AAU26007	Aau26007	Human thr	486	31	91.2	18	5	ABP51673	Abp51673	TPO mimet
414	31	91.2	15	4	AAU26038	Aau26038	Human thr	487	31	91.2	18	5	ABP51690	Abp51690	TPO mimet
415	31	91.2	15	5	ABP51670	Abp51670	Thrombopo	488	31	91.2	18	5	ABP51675	Abp51675	TPO mimet
416	31	91.2	15	7	ABR62908	Abmr62908	Thrombopo	489	31	91.2	18	5	ABP51692	Abp51692	TPO mimet
417	31	91.2	15	8	ADM72485	Adm72485	TPO mimet	490	31	91.2	18	5	ABP51679	Abp51679	TPO mimet
418	31	91.2	15	8	ADM72479	Adm72479	TPO mimet	491	31	91.2	18	7	ADJ73062	Adj73062	TPO mimet
419	31	91.2	15	8	ADM72502	Adm72502	TPO mimet	492	31	91.2	18	7	ADJ73063	Adj73063	TPO mimet
420	31	91.2	15	8	ADM72492	Adm72492	TPO mimet	493	31	91.2	18	7	ADN59681	Adn59681	Thrombopo
421	31	91.2	15	8	ADM72478	Adm72478	TPO mimet	494	31	91.2	18	7	ADN59660	Adn59660	Thrombopo
422	31	91.2	15	8	ADM72533	Adm72533	TPO mimet	495	31	91.2	18	7	ADN59663	Adn59663	Thrombopo
423	31	91.2	15	8	ADM72496	Adm72496	TPO mimet	496	31	91.2	18	7	ADN59679	Adn59679	Thrombopo
424	31	91.2	15	8	ADM72490	Adm72490	TPO mimet	497	31	91.2	18	7	ADN59654	Adn59654	Thrombopo
425	31	91.2	15	8	ADM72486	Adm72486	TPO mimet	498	31	91.2	18	7	ADN59678	Adn59678	Thrombopo
426	31	91.2	15	8	ADM72491	Adm72491	TPO mimet	499	31	91.2	18	7	ADN59815	Adn59815	Thrombopo
427	31	91.2	15	8	ADM72522	Adm72522	TPO mimet	500	31	91.2	18	7	ADN59661	Adn59661	Thrombopo
428	31	91.2	15	8	ADM72523	Adm72523	TPO mimet	501	31	91.2	18	7	ADN59653	Adn59653	Thrombopo
429	31	91.2	15	8	ADM72493	Adm72493	TPO mimet	502	31	91.2	18	7	ADN59676	Adn59676	Thrombopo
430	31	91.2	15	8	ADM72482	Adm72482	TPO mimet	503	31	91.2	18	7	ADN59812	Adn59812	Thrombopo
431	31	91.2	15	8	ADM72494	Adm72494	TPO mimet	504	31	91.2	18	7	ADN59666	Adn59666	Thrombopo
432	31	91.2	15	8	ADQ16585	Adq16585	TPO mimet	505	31	91.2	18	7	ADN59670	Adn59670	Thrombopo
433	31	91.2	15	8	ADU92483	Adt92483	Modified	506	31	91.2	18	7	ADN59656	Adn59656	Thrombopo
434	31	91.2	15	9	ADU70209	Adu70209	Thrombopo	507	31	91.2	18	8	ADJ52698	Adj52698	CHI delet
435	31	91.2	15	9	ADU75981	Adu75981	Peptide-b	508	31	91.2	18	8	ADJ52697	Adj52697	CHI delet
436	31	91.2	15	9	ADV44320	Adv44320	Agonist T	509	31	91.2	18	8	ADJ51658	Adj51658	CHI delet
437	31	91.2	15	9	ABE12793	Aeb12793	TPO mimet	510	31	91.2	18	8	ADJ51659	Adj51659	CHI delet
438	31	91.2	16	2	AAW19534	Aaw19534	Thrombopo	511	31	91.2	18	8	ADQ16611	Adq16611	TPO mimet
439	31	91.2	16	2	AAW33035	Aaw33035	Thrombopo	512	31	91.2	18	8	ADQ16619	Adq16619	TPO mimet
440	31	91.2	16	2	AAW36775	Aaw36775	Thrombopo	513	31	91.2	18	8	ADQ16621	Adq16621	TPO mimet
441	31	91.2	16	2	AAW36771	Aaw36771	Thrombopo	514	31	91.2	18	8	ADQ16641	Adq16641	TPO mimet
442	31	91.2	16	2	AAW66709	Aaw66709	Peptide c	515	31	91.2	18	8	ADQ16646	Adq16646	TPO mimet
443	31	91.2	16	2	AAW66713	Aaw66713	Peptide c	516	31	91.2	18	8	ADQ16607	Adq16607	TPO mimet
444	31	91.2	16	2	AAW66733	Aaw66733	Peptide c	517	31	91.2	18	8	ADQ16615	Adq16615	TPO mimet
445	31	91.2	16	2	AAW66716	Aaw66716	Peptide c	518	31	91.2	18	8	ADQ16693	Adq16693	TPO mimet
446	31	91.2	16	4	AAU26021	Aau26021	Human thr	519	31	91.2	18	8	ADQ16627	Adq16627	TPO mimet
447	31	91.2	16	4	AAU26005	Aau26005	Human thr	520	31	91.2	18	8	ADQ16625	Adq16625	TPO mimet
448	31	91.2	16	4	AAU26043	Aau26043	Human thr	521	31	91.2	18	8	ADQ16617	Adq16617	TPO mimet
449	31	91.2	16	4	AAU25832	Aau25832	Human thr	522	31	91.2	18	8	ADQ16629	Adq16629	TPO mimet
450	31	91.2	16	8	ADM72532	Adm72532	TPO mimet	523	31	91.2	18	8	ADQ16613	Adq16613	TPO mimet
451	31	91.2	16	8	ADM72484	Adm72484	TPO mimet	524	31	91.2	18	8	ADQ16623	Adq16623	TPO mimet
452	31	91.2	18	2	AAW09592	Aaw09592	Thrombopo	525	31	91.2	18	8	ADQ16605	Adq16605	TPO mimet
453	31	91.2	18	2	AAW09499	Aaw09499	Thrombopo	526	31	91.2	18	8	ADQ16609	Adq16609	TPO mimet
454	31	91.2	18	2	AAW09595	Aaw09595	Thrombopo	527	31	91.2	18	9	ADV44355	Adv44355	Agonist T
455	31	91.2	18	2	AAW09497	Aaw09497	Thrombopo	528	31	91.2	18	9	ADV44431	Adv44431	Anti-teta
456	31	91.2	18	2	AAW09459	Aaw09459	Thrombopo	529	31	91.2	18	9	ADV44345	Adv44345	Agonist T
457	31	91.2	18	2	AAW36650	Aaw36650	Thrombopo	530	31	91.2	18	9	ADV44351	Adv44351	Agonist T
458	31	91.2	18	2	AAW36746	Aaw36746	Thrombopo	531	31	91.2	18	9	ADV44357	Adv44357	Agonist T
459	31	91.2	18	2	AAW36653	Aaw36653	Thrombopo	532	31	91.2	18	9	ADV44349	Adv44349	Agonist T
460	31	91.2	18	2	AAW33026	Aaw33026	Thrombopo	533	31	91.2	18	9	ADV44343	Adv44343	Agonist T
461	31	91.2	18	2	AAW36743	Aaw36743	Thrombopo	534	31	91.2	18	9	ADV44347	Adv44347	Agonist T
462	31	91.2	18	2	AAW36648	Aaw36648	Thrombopo	535	31	91.2	18	9	ADV44384	Adv44384	Agonist T

536	31	91.2	18	9	ADV44353	Adv44353 Agonist T	609	31	91.2	22	8	ADQ16714	Adq16714 Immunoglo
537	31	91.2	18	9	ADV44361	Adv44363 Agonist T	610	31	91.2	22	8	ADQ16713	Adq16713 Immunoglo
538	31	91.2	18	9	ADV44361	Adv44361 Agonist T	611	31	91.2	22	8	ADQ16709	Adq16709 Immunoglo
539	31	91.2	18	9	ADV44379	Adv44379 Agonist T	612	31	91.2	22	8	ADQ16706	Adq16706 Immunoglo
540	31	91.2	18	9	ADV44344	Adv44344 Agonist T	613	31	91.2	22	8	ADQ16699	TPO mimet
541	31	91.2	18	9	ADV44359	Adv44359 Agonist T	614	31	91.2	22	8	ADQ16712	Immunoglo
542	31	91.2	18	9	ADV44365	Adv44365 Agonist T	615	31	91.2	22	8	ADQ16707	Immunoglo
543	31	91.2	18	9	ADV44367	Adv44367 Agonist T	616	31	91.2	22	8	ADQ16711	Immunoglo
544	31	91.2	18	9	ADV44468	Adv44468 Agonist T	617	31	91.2	22	8	ADQ16708	Immunoglo
545	31	91.2	18	9	ABE12826	Abel2826 TPO mimet	618	31	91.2	22	8	ADQ16710	Immunoglo
546	31	91.2	18	9	ABE12820	Abel2820 TPO mimet	619	31	91.2	22	9	ADV44435	Modified
547	31	91.2	18	9	ABE12840	Abel2840 TPO mimet	620	31	91.2	22	9	ADV44449	Anti-teta
548	31	91.2	18	9	ABE12857	Abel2857 Antibody	621	31	91.2	22	9	ADV44443	Anti-teta
549	31	91.2	18	9	ABE12834	Abel2834 TPO mimet	622	31	91.2	22	9	ADV44444	Anti-teta
550	31	91.2	18	9	ABE12939	Abel2939 TPO mimet	623	31	91.2	22	9	ADV44448	Anti-teta
551	31	91.2	18	9	ABE12832	Abel2832 TPO mimet	624	31	91.2	22	9	ADV44442	Anti-teta
552	31	91.2	18	9	ABE12852	Abel2852 TPO mimet	625	31	91.2	22	9	ADV44447	Anti-teta
553	31	91.2	18	9	ABE12816	Abel2816 TPO mimet	626	31	91.2	22	9	ADV44445	Anti-teta
554	31	91.2	18	9	ABE12818	Abel2818 TPO mimet	627	31	91.2	22	9	ADV44446	Anti-teta
555	31	91.2	18	9	ABE12828	Abel2828 TPO mimet	628	31	91.2	22	9	ABE12918	TPO mimet
556	31	91.2	18	9	ABE12836	Abel2836 TPO mimet	629	31	91.2	22	9	ABE12917	TPO mimet
557	31	91.2	18	9	ABE12836	Abel2836 TPO mimet	630	31	91.2	22	9	ABE12921	TPO mimet
558	31	91.2	18	9	ABE12903	Abel2903 Heavy cha	631	31	91.2	22	9	ABE12919	Tt antibo
559	31	91.2	18	9	ABE12822	Abel2822 TPO mimet	632	31	91.2	22	9	ABE12986	Tt antibo
560	31	91.2	18	9	ABE12824	Abel2824 TPO mimet	633	31	91.2	22	9	ABE12920	TPO mimet
561	31	91.2	18	9	ABE12838	Abel2838 TPO mimet	634	31	91.2	22	9	ABE12914	TPO mimet
562	31	91.2	19	2	AAW09494	Aaw09494 Thrombopo	635	31	91.2	22	9	ABE12921	TPO mimet
563	31	91.2	19	2	AAW09461	Aaw09461 Thrombopo	636	31	91.2	22	9	ABE12907	Tt antibo
564	31	91.2	19	2	AAW09491	Aaw09491 Thrombopo	637	31	91.2	22	9	ABE12915	TPO mimet
565	31	91.2	19	2	AAW09493	Aaw09493 Thrombopo	638	31	91.2	22	9	ABE12918	TPO mimet
566	31	91.2	19	2	AAW09457	Aaw09457 Thrombopo	639	31	91.2	25	7	ADN59736	Thrombopo
567	31	91.2	19	2	AAW09492	Aaw09492 Thrombopo	640	31	91.2	25	7	ADN59742	Thrombopo
568	31	91.2	19	2	AAW36651	Aaw36651 Thrombopo	641	31	91.2	25	7	ADN59689	Thrombopo
569	31	91.2	19	2	AAW33028	Aaw33028 Thrombopo	642	31	91.2	25	7	ADN59704	Thrombopo
570	31	91.2	19	2	AAW33024	Aaw33024 Thrombopo	643	31	91.2	25	7	ADN59732	Thrombopo
571	31	91.2	19	2	AAW36643	Aaw36643 Thrombopo	644	31	91.2	25	7	ADN59708	Thrombopo
572	31	91.2	19	2	AAW36645	Aaw36645 Thrombopo	645	31	91.2	25	7	ADN59732	Thrombopo
573	31	91.2	19	2	AAW36645	Aaw36645 Thrombopo	646	31	91.2	25	7	ADN59738	Thrombopo
574	31	91.2	19	2	AAW35418	Aaw35418 Thrombopo	647	31	91.2	25	7	ADN59702	Thrombopo
575	31	91.2	19	2	AAW36642	Aaw36642 Thrombopo	648	31	91.2	25	7	ADN59691	Thrombopo
576	31	91.2	19	3	ABE17022	Abel17022 TPO-mimet	649	31	91.2	25	7	ADN59708	Thrombopo
577	31	91.2	19	3	ABE17021	Abel17021 TPO-mimet	650	31	91.2	28	3	AAE17285	TPO-mimet
578	31	91.2	19	4	AAU25861	Aau25861 Human thr	651	31	91.2	28	5	ABP51682	TPO mimet
579	31	91.2	19	4	AAU25864	Aau25864 Human thr	652	31	91.2	28	8	ADJ73013	TPO mimet
580	31	91.2	19	4	AAU25864	Aau25864 Human thr	653	31	91.2	28	8	ADJ52648	CH1 delet
581	31	91.2	19	4	AAU25862	Aau25862 Human thr	654	31	91.2	28	8	ADJ51609	CH1 delet
582	31	91.2	19	4	AAU25870	Aau25870 Human thr	655	31	91.2	28	8	ADQ16636	Tetanus t
583	31	91.2	19	4	AAU25825	Aau25825 Human thr	656	31	91.2	28	8	ADV44374	Modified
584	31	91.2	19	4	AAU25998	Aau25998 Human thr	657	31	91.2	28	9	ABE12847	Antibody
585	31	91.2	19	4	AAU25821	Aau25821 Human thr	658	31	91.2	28	9	ABE16971	TPO-mimet
586	31	91.2	19	5	ABB73391	Abb73391 TPO-mimet	659	31	91.2	29	3	AAE16975	TPO-mimet
587	31	91.2	19	5	ABB72908	Abb72908 TPO mimet	660	31	91.2	29	3	AAE16976	TPO-mimet
588	31	91.2	19	5	ABB72907	Abb72907 TPO mimet	661	31	91.2	29	3	AAE17286	TPO-mimet
589	31	91.2	19	5	ABB73390	Abb73390 TPO-mimet	662	31	91.2	29	3	AAE16970	TPO-mimet
590	31	91.2	19	7	ADJ73059	Adj73059 TPO mimet	663	31	91.2	29	5	ABE72862	TPO-mimet
591	31	91.2	19	8	ADJ73059	Adj73059 TPO mimet	664	31	91.2	29	5	ABE72857	TPO mimet
592	31	91.2	19	8	ADJ52695	Adj52695 CH1 delet	665	31	91.2	29	5	ABE72861	TPO mimet
593	31	91.2	19	8	ADJ52694	Adj52694 CH1 delet	666	31	91.2	29	5	ABE72856	TPO mimet
594	31	91.2	19	8	ADJ51655	Adj51655 CH1 delet	667	31	91.2	29	7	ADJ73011	TPO mimet
595	31	91.2	19	8	ADJ51656	Adj51656 CH1 delet	668	31	91.2	29	7	ADJ73007	TPO mimet
596	31	91.2	20	3	AAE18003	Aae18003 FC-TMP de	669	31	91.2	29	8	ADJ73006	TPO mimet
597	31	91.2	20	3	AAE17929	Aae17929 TPO-mimet	670	31	91.2	29	8	ADJ52642	CH1 delet
598	31	91.2	20	5	ABB73403	Abb73403 TPO mimet	671	31	91.2	29	8	ADJ52646	CH1 delet
599	31	91.2	21	7	ADN59687	Adn59687 Thrombopo	672	31	91.2	29	8	ADJ52641	CH1 delet
600	31	91.2	22	7	ADN59820	Adn59820 TWP pepti	673	31	91.2	29	8	ADJ51603	CH1 delet
601	31	91.2	22	7	ADN59821	Adn59821 TWP pepti	674	31	91.2	29	8	ADJ51602	CH1 delet
602	31	91.2	22	7	ADN59828	Adn59828 TWP pepti	675	31	91.2	29	8	ADJ51607	CH1 delet
603	31	91.2	22	7	ADN59830	Adn59830 TWP pepti	676	31	91.2	30	3	AAE17287	TPO-mimet
604	31	91.2	22	7	ADN59823	Adn59823 TWP pepti	677	31	91.2	31	3	AAE17288	TPO-mimet
605	31	91.2	22	7	ADN59827	Adn59827 TWP pepti	678	31	91.2	31	3	AAE16974	TPO-mimet
606	31	91.2	22	7	ADN59833	Adn59833 TWP pepti	679	31	91.2	31	3	AAE16973	TPO-mimet
607	31	91.2	22	7	ADN59819	Adn59819 TWP pepti	680	31	91.2	31	5	ABE72860	TPO mimet
608	31	91.2	22	7	ADN59837	Adn59837 TWP pepti	681	31	91.2	31			

682	31	91.2	31	5	ABB72859	Abb72859 TPO minet	755	31	91.2	143	6	ABG71750	Abg71750 Antibody
683	31	91.2	31	7	ADJ73009	Adj73009 TPO minet	756	31	91.2	144	6	ABG71748	Abg71748 Antibody
684	31	91.2	31	7	ADJ73010	Adj73010 TPO minet	757	31	91.2	150	9	ABE12945	AbE12945 Antibody
685	31	91.2	31	8	ADJ52644	Adj52644 CH1 delet	758	31	91.2	158	4	AAU29357	Aau29357 Novel mar
686	31	91.2	31	8	ADJ52645	Adj52645 CH1 delet	759	31	91.2	165	9	ABM92612	Abm92612 M. xanthu
687	31	91.2	31	8	ADJ51606	Adj51606 CH1 delet	760	31	91.2	225	8	ADQ16704	Adq16704 Modified
688	31	91.2	31	8	ADJ51605	Adj51605 CH1 delet	761	31	91.2	234	9	ABE12912	AbE12912 Antibody
689	31	91.2	32	3	AAV96520	Aav96520 Thrombopo	762	31	91.2	247	3	AAE16958	Aae16958 FC-TMP pr
690	31	91.2	32	3	AAE17289	Aae17289 TPO-minet	763	31	91.2	247	3	ABE16961	AbE16961 TMP-FC pr
691	31	91.2	32	3	AAE17297	Aae17297 TPO-minet	764	31	91.2	247	5	ABB73411	Abb73411 FC-TPO mi
692	31	91.2	32	3	AAE17290	Aae17290 TPO-minet	765	31	91.2	247	5	ABB73414	Abb73414 TMP-FC am
693	31	91.2	34	3	AAV96527	Aav96527 Thrombopo	766	31	91.2	249	9	ADV44440	Adv44440 pAX116 va
694	31	91.2	34	3	AAV96527	Aav96527 Thrombopo	767	31	91.2	268	3	AAE16959	Aae16959 FC-TMP-TM
695	31	91.2	35	3	ABE17291	AbE17291 TPO-minet	768	31	91.2	268	5	ABE16958	AbE16958 FC-TMP-TM
696	31	91.2	35	3	ABE17292	AbE17292 TPO-minet	769	31	91.2	269	3	AAV96531	Aav96531 Human Igg
697	31	91.2	36	3	AAV96525	Aav96525 Thrombopo	770	31	91.2	269	3	ABE16960	AbE16960 TMP-TMP-F
698	31	91.2	36	3	AAV96524	Aav96524 Thrombopo	771	31	91.2	269	5	ABB73413	Abb73413 TMP-TMP-F
699	31	91.2	36	3	AAV96526	Aav96526 Thrombopo	772	31	91.2	282	8	ADS24597	AdS24597 Bacterial
700	31	91.2	36	3	AAE17307	Aae17307 TPO-minet	773	31	91.2	282	9	ABE12930	AbE12930 Antibody
701	31	91.2	36	3	ABE17293	AbE17293 TPO-minet	774	31	91.2	303	4	ABU53220	Abu53220 Human met
702	31	91.2	36	3	ABE17303	AbE17303 TPO-minet	775	31	91.2	308	4	RAG91969	Rag91969 C glucami
703	31	91.2	36	3	AAE16963	Aae16963 TPO-minet	776	31	91.2	376	8	ABM84031	Abm84031 Human dia
704	31	91.2	36	3	AAE17301	Aae17301 TPO-minet	777	31	91.2	377	8	ABM84030	Abm84030 Human dia
705	31	91.2	36	3	AAE17306	Aae17306 TPO-minet	778	31	91.2	404	6	ABP56521	Abp56521 Pseudomon
706	31	91.2	36	5	ABB72403	Abb72403 TPO-minet	779	31	91.2	405	2	AAV04128	Aav04128 Pseudomon
707	31	91.2	37	3	ABE17294	AbE17294 TPO-minet	780	31	91.2	410	6	ADA66611	Ada66611 Mouse Tra
708	31	91.2	38	3	ABE17295	AbE17295 TPO-minet	781	31	91.2	412	2	AAE08264	Aae08264 Tumour gr
709	31	91.2	39	3	AAE17304	Aae17304 TPO-minet	782	31	91.2	412	2	AAE08261	Aae08261 Transform
710	31	91.2	39	3	AAE17305	Aae17305 TPO-minet	783	31	91.2	412	2	AAV73598	Aav73598 Human TGF
711	31	91.2	40	3	AAE17302	Aae17302 TPO-minet	784	31	91.2	412	2	AAW80417	Aaw80417 Amino aci
712	31	91.2	40	7	ADN59753	Adn59753 Peptide-v	785	31	91.2	412	5	ABB90765	Abb90765 Human Tum
713	31	91.2	40	7	ADN59760	Adn59760 Peptide-v	786	31	91.2	412	5	AAE29083	Aae29083 Human tra
714	31	91.2	41	3	AAV96528	Aav96528 Thrombopo	787	31	91.2	412	5	AAU77103	Aau77103 Human tra
715	31	91.2	41	5	ABB73389	Abb73389 TPO-minet	788	31	91.2	412	6	ADA66610	Ada66610 Human tra
716	31	91.2	41	5	ABB73388	Abb73388 TPO-minet	789	31	91.2	412	6	ABU54472	Abu54472 Human tum
717	31	91.2	42	3	AAV96530	Aav96530 Thrombopo	790	31	91.2	412	8	ADH11592	Adh11592 Human bon
718	31	91.2	42	3	ABE17296	AbE17296 TPO-minet	791	31	91.2	412	8	ADS87980	Ads87980 Tumour tr
719	31	91.2	42	3	ABE17308	AbE17308 Synthetic	792	31	91.2	412	8	ADU18086	Adu18086 Human can
720	31	91.2	42	3	AAE17282	Aae17282 TPO-minet	793	31	91.2	412	9	ADY27744	Ady27744 Human tra
721	31	91.2	42	3	AAE17281	Aae17281 TPO-minet	794	31	91.2	412	9	ABE01311	AbE01311 Human TGF
722	31	91.2	42	5	ABB73404	Abb73404 TMP-TMP g	795	31	91.2	412	9	ABE44146	AbE44146 Human tra
723	31	91.2	42	7	ADN59751	Adn59751 Peptide-v	796	31	91.2	413	2	AAE22038	Aae22038 Mutant tr
724	31	91.2	43	7	ADN59752	Adn59752 Peptide-v	797	31	91.2	414	3	AAV33289	Aav33289 Pseudomon
725	31	91.2	43	7	ADN59759	Adn59759 Peptide-v	798	31	91.2	414	3	AAV93565	Aav93565 Amino aci
726	31	91.2	44	7	ADN59817	Adn59817 Peptide-v	799	31	91.2	414	4	AAV72225	Aav72225 Pseudomon
727	31	91.2	60	3	AAE17311	Aae17311 Synthetic	800	31	91.2	414	4	AAV72237	Aav72237 Pseudomon
728	31	91.2	60	5	ABE73405	AbE73405 TMP-TMP g	801	31	91.2	414	4	AAV72245	Aav72245 Pseudomon
729	31	91.2	68	4	ABE39037	AbE39037 Peptide #	802	31	91.2	414	4	AAV72246	Aav72246 Pseudomon
730	31	91.2	68	4	AAE32526	Aae32526 Peptide #	803	31	91.2	414	4	AAV72265	Aav72265 Pseudomon
731	31	91.2	68	4	AAV72267	Aav72267 Human liv	804	31	91.2	414	4	AAV72272	Aav72272 Pseudomon
732	31	91.2	68	4	ABG53953	Abg53953 Human liv	805	31	91.2	414	4	AAV72279	Aav72279 Pseudomon
733	31	91.2	68	5	ABG42082	Abg42082 Human pep	806	31	91.2	414	4	AAV72291	Aav72291 Pseudomon
734	31	91.2	114	3	AAE25241	Aae25241 Eucalyptu	807	31	91.2	414	4	AAV72300	Aav72300 Pseudomon
735	31	91.2	122	9	ADV44474	Adv44474 Anti-teta	808	31	91.2	414	4	AAV72301	Aav72301 Pseudomon
736	31	91.2	122	9	ABE12946	AbE12946 Antibody	809	31	91.2	414	4	AAV72336	Aav72336 P. putida
737	31	91.2	128	8	ADQ16705	Adq16705 Modified	810	31	91.2	414	4	AAV72213	Aav72213 Pseudomon
738	31	91.2	128	8	ADV44466	Adv44466 Anti-teta	811	31	91.2	414	4	AAV72214	Aav72214 Pseudomon
739	31	91.2	128	9	ADV44463	Adv44463 Anti-teta	812	31	91.2	414	4	AAV72220	Aav72220 Pseudomon
740	31	91.2	128	9	ADV44467	Adv44467 Anti-teta	813	31	91.2	414	4	AAV72226	Aav72226 Pseudomon
741	31	91.2	128	9	ADV44465	Adv44465 Anti-teta	814	31	91.2	414	4	AAV72231	Aav72231 Pseudomon
742	31	91.2	128	9	ADV44441	Adv44441 pAX116 va	815	31	91.2	414	4	AAV72234	Aav72234 Pseudomon
743	31	91.2	128	9	ADV44464	Adv44464 Anti-teta	816	31	91.2	414	4	AAV72240	Aav72240 Pseudomon
744	31	91.2	128	9	ABE12934	AbE12934 Antibody	817	31	91.2	414	4	AAV72244	Aav72244 Pseudomon
745	31	91.2	128	9	ABE12935	AbE12935 Antibody	818	31	91.2	414	4	AAV72247	Aav72247 Pseudomon
746	31	91.2	128	9	ABE12913	AbE12913 Antibody	819	31	91.2	414	4	AAV72266	Aav72266 Pseudomon
747	31	91.2	128	9	ABE12936	AbE12936 Antibody	820	31	91.2	414	4	AAV72298	Aav72298 Pseudomon
748	31	91.2	128	9	ABE12937	AbE12937 Antibody	821	31	91.2	414	4	AAV72319	Aav72319 Pseudomon
749	31	91.2	128	9	ABE12938	AbE12938 Antibody	822	31	91.2	414	4	AAV72338	Aav72338 P. putida
750	31	91.2	129	6	ABG71751	Abg71751 Antibody	823	31	91.2	414	4	AAV72255	Aav72255 Pseudomon
751	31	91.2	131	6	ABG71753	Abg71753 Antibody	824	31	91.2	414	4	AAV72270	Aav72270 Pseudomon
752	31	91.2	132	9	ADV44473	Adv44473 Anti-teta	825	31	91.2	414	4	AAV72273	Aav72273 Pseudomon
753	31	91.2	133	6	ABG71752	Abg71752 Antibody	826	31	91.2	414	4	AAV72274	Aav72274 Pseudomon
754	31	91.2	135	6	ABG71749	Abg71749 Antibody	827	31	91.2	414	4	AAV72275	Aav72275 Pseudomon

828	31	91.2	414	4	AAY72306	Aay72306	Pseudomon	901	31	91.2	414	4	AAY72314	Aay72314	Pseudomon
829	31	91.2	414	4	AAY72309	Aay72309	Pseudomon	902	31	91.2	414	4	AAY72318	Aay72318	Pseudomon
830	31	91.2	414	4	AAY72312	Aay72312	Pseudomon	903	31	91.2	414	4	AAY72327	Aay72327	Pseudomon
831	31	91.2	414	4	AAY72331	Aay72331	Pseudomon	904	31	91.2	414	4	AAY72229	Aay72229	Pseudomon
832	31	91.2	414	4	AAY72233	Aay72233	Pseudomon	905	31	91.2	414	4	AAY72242	Aay72242	Pseudomon
833	31	91.2	414	4	AAY72236	Aay72236	Pseudomon	906	31	91.2	414	4	AAY72250	Aay72250	Pseudomon
834	31	91.2	414	4	AAY72260	Aay72260	Pseudomon	907	31	91.2	414	4	AAY72285	Aay72285	Pseudomon
835	31	91.2	414	4	AAY72261	Aay72261	Pseudomon	908	31	91.2	414	4	AAY72289	Aay72289	Pseudomon
836	31	91.2	414	4	AAY72276	Aay72276	Pseudomon	909	31	91.2	414	4	AAY72295	Aay72295	Pseudomon
837	31	91.2	414	4	AAY72278	Aay72278	Pseudomon	910	31	91.2	414	4	AAY72299	Aay72299	Pseudomon
838	31	91.2	414	4	AAY72305	Aay72305	Pseudomon	911	31	91.2	414	4	AAY72320	Aay72320	Pseudomon
839	31	91.2	414	4	AAY72317	Aay72317	Pseudomon	912	31	91.2	414	4	AAY72335	Aay72335	Pseudomon
840	31	91.2	414	4	AAY72324	Aay72324	Pseudomon	913	31	91.2	414	4	AAY72337	Aay72337	P. putida
841	31	91.2	414	4	AAY72328	Aay72328	Pseudomon	914	31	91.2	414	4	AAY72217	Aay72217	Pseudomon
842	31	91.2	414	4	AAY72342	Aay72342	Pseudomon	915	31	91.2	414	4	AAY72253	Aay72253	Pseudomon
843	31	91.2	414	4	AAY72241	Aay72241	Pseudomon	916	31	91.2	414	4	AAY72269	Aay72269	Pseudomon
844	31	91.2	414	4	AAY72257	Aay72257	Pseudomon	917	31	91.2	414	4	AAY72303	Aay72303	Pseudomon
845	31	91.2	414	4	AAY72287	Aay72287	Pseudomon	918	31	91.2	414	4	AAY72307	Aay72307	Pseudomon
846	31	91.2	414	4	AAY72311	Aay72311	Pseudomon	919	31	91.2	414	4	AAY72224	Aay72224	Pseudomon
847	31	91.2	414	4	AAY72321	Aay72321	Pseudomon	920	31	91.2	414	4	AAY72239	Aay72239	Pseudomon
848	31	91.2	414	4	AAY72207	Aay72207	Pseudomon	921	31	91.2	414	4	AAY72268	Aay72268	Pseudomon
849	31	91.2	414	4	AAY72219	Aay72219	Pseudomon	922	31	91.2	414	4	AAY72280	Aay72280	Pseudomon
850	31	91.2	414	4	AAY72221	Aay72221	Pseudomon	923	31	91.2	414	4	AAY72296	Aay72296	Pseudomon
851	31	91.2	414	4	AAY72235	Aay72235	Pseudomon	924	31	91.2	414	4	AAY72302	Aay72302	Pseudomon
852	31	91.2	414	4	AAY72254	Aay72254	Pseudomon	925	31	91.2	414	4	AAY72329	Aay72329	Pseudomon
853	31	91.2	414	4	AAY72263	Aay72263	Pseudomon	926	31	91.2	414	4	AAY72330	Aay72330	Pseudomon
854	31	91.2	414	4	AAY72281	Aay72281	Pseudomon	927	31	91.2	414	4	AAY72332	Aay72332	Pseudomon
855	31	91.2	414	4	AAY72292	Aay72292	Pseudomon	928	31	91.2	414	5	ABO8123	ABO8123	P. putida
856	31	91.2	414	4	AAY72310	Aay72310	Pseudomon	929	31	91.2	414	5	ABO8127	ABO8127	P. putida
857	31	91.2	414	4	AAY72316	Aay72316	Pseudomon	930	31	91.2	414	5	ABO8126	ABO8126	P. putida
858	31	91.2	414	4	AAY72227	Aay72227	Pseudomon	931	31	91.2	414	5	ABO8128	ABO8128	P. putida
859	31	91.2	414	4	AAY72248	Aay72248	Pseudomon	932	31	91.2	414	6	ABR43996	ABR43996	P. putida
860	31	91.2	414	4	AAY72282	Aay72282	Pseudomon	933	31	91.2	414	8	ADO47232	ADO47232	Proteomase-
861	31	91.2	414	4	AAY72283	Aay72283	Pseudomon	934	31	91.2	414	9	ADV90218	ADV90218	Pseudomon
862	31	91.2	414	4	AAY72326	Aay72326	Pseudomon	935	31	91.2	414	9	ADV69262	ADV69262	P. putida
863	31	91.2	414	4	AAY72215	Aay72215	Pseudomon	936	31	91.2	423	8	ADQ17024	ADQ17024	Murine TG
864	31	91.2	414	4	AAY72249	Aay72249	Pseudomon	937	31	91.2	427	8	ADQ17026	ADQ17026	Pig TGF-b
865	31	91.2	414	4	AAY72284	Aay72284	Pseudomon	938	31	91.2	456	2	AAW78786	AAW78786	Pig trans
866	31	91.2	414	4	AAY72308	Aay72308	Pseudomon	939	31	91.2	456	2	AAW78786	AAW78786	Pig trans
867	31	91.2	414	4	AAY72347	Aay72347	Pseudomon	940	31	91.2	457	2	AAW44459	AAW44459	Anti-t-teta
868	31	91.2	414	4	AAY72215	Aay72215	Pseudomon	941	31	91.2	459	9	ADP51695	ADP51695	SG1-1-TPO
869	31	91.2	414	4	AAY72238	Aay72238	Pseudomon	942	31	91.2	472	8	ADQ16647	ADQ16647	Immunoglo
870	31	91.2	414	4	AAY72243	Aay72243	Pseudomon	943	31	91.2	472	8	ADQ16647	ADQ16647	Immunoglo
871	31	91.2	414	4	AAY72251	Aay72251	Pseudomon	944	31	91.2	472	8	ADQ16647	ADQ16647	Immunoglo
872	31	91.2	414	4	AAY72264	Aay72264	Pseudomon	945	31	91.2	472	8	ADQ16647	ADQ16647	Immunoglo
873	31	91.2	414	4	AAY72293	Aay72293	Pseudomon	946	31	91.2	531	6	ABU29925	ABU29925	Protein e
874	31	91.2	414	4	AAY72323	Aay72323	Pseudomon	947	31	91.2	531	6	ABU29925	ABU29925	Protein e
875	31	91.2	414	4	AAY72256	Aay72256	Pseudomon	948	31	91.2	543	7	ADC94090	ADC94090	E. faeciu
876	31	91.2	414	4	AAY72258	Aay72258	Pseudomon	949	31	91.2	602	7	ABO66004	ABO66004	Krebiell
877	31	91.2	414	4	AAY72262	Aay72262	Pseudomon	950	31	91.2	637	7	ABO66004	ABO66004	Krebiell
878	31	91.2	414	4	AAY72271	Aay72271	Pseudomon	951	31	91.2	653	8	ADS30548	ADS30548	Bacterial
879	31	91.2	414	4	AAY72277	Aay72277	Pseudomon	952	31	91.2	744	9	AE840488	AE840488	L. pneumo
880	31	91.2	414	4	AAY72334	Aay72334	Pseudomon	953	31	91.2	751	9	AE837147	AE837147	L. pneumo
881	31	91.2	414	4	AAY72218	Aay72218	Pseudomon	954	31	91.2	984	5	ABU15501	ABU15501	Herbicida
882	31	91.2	414	4	AAY72222	Aay72222	Pseudomon	955	31	91.2	1043	6	ABU15501	ABU15501	Herbicida
883	31	91.2	414	4	AAY72228	Aay72228	Pseudomon	956	31	91.2	1094	4	ABU15501	ABU15501	Herbicida
884	31	91.2	414	4	AAY72230	Aay72230	Pseudomon	957	31	91.2	1151	7	ABO74850	ABO74850	Protein e
885	31	91.2	414	4	AAY72257	Aay72257	Pseudomon	958	31	91.2	1220	6	ABU41578	ABU41578	Protein e
886	31	91.2	414	4	AAY72304	Aay72304	Pseudomon	959	31	91.2	1630	6	ABU44550	ABU44550	Human wou
887	31	91.2	414	4	AAY72304	Aay72304	Pseudomon	960	31	91.2	1630	6	ABU44550	ABU44550	Human wou
888	31	91.2	414	4	AAY72333	Aay72333	Pseudomon	961	31	91.2	1632	4	ABU44550	ABU44550	Human wou
889	31	91.2	414	4	AAY72267	Aay72267	Pseudomon	962	31	91.2	2955	4	ABG21509	ABG21509	Novel hum
890	31	91.2	414	4	AAY72286	Aay72286	Pseudomon	963	31	91.2	2988	4	ABG21509	ABG21509	Novel hum
891	31	91.2	414	4	AAY72288	Aay72288	Pseudomon	964	31	91.2	2988	4	ABG21509	ABG21509	Novel hum
892	31	91.2	414	4	AAY72290	Aay72290	Pseudomon	965	31	91.2	4342	6	ABU15500	ABU15500	Protein e
893	31	91.2	414	4	AAY72313	Aay72313	Pseudomon	966	31	91.2	4342	6	ABU15500	ABU15500	Protein e
894	31	91.2	414	4	AAY72315	Aay72315	Pseudomon	967	31	91.2	88.2	12	AAW36733	AAW36733	Thrombopo
895	31	91.2	414	4	AAY72316	Aay72316	Pseudomon	968	31	91.2	88.2	12	AAW36733	AAW36733	Thrombopo
896	31	91.2	414	4	AAY72325	Aay72325	Pseudomon	969	31	91.2	88.2	12	AAW36733	AAW36733	Thrombopo
897	31	91.2	414	4	AAY72212	Aay72212	Pseudomon	970	31	91.2	88.2	16	AAW36760	AAW36760	Thrombopo
898	31	91.2	414	4	AAY72259	Aay72259	Pseudomon	971	31	91.2	88.2	16	AAW36760	AAW36760	Thrombopo
899	31	91.2	414	4	AAY72294	Aay72294	Pseudomon	972	31	91.2	88.2	18	AAW00367	AAW00367	Streptomy
900	31	91.2	414	4	AAY72297	Aay72297	Pseudomon	973	31	91.2	88.2	19	AAW09495	AAW09495	Thrombopo

974 30 88.2 19 2 AAW36646 Thrombopo
975 30 88.2 19 3 AAB17023 TPO-mimet
976 30 88.2 19 4 AAU25865 Human thr
977 30 88.2 19 5 ABB72909 TPO mimet
978 30 88.2 19 7 ADJ73061 TPO mimet
979 30 88.2 19 8 ADJ52696 CH1 delet
980 30 88.2 19 8 ADJ51657 CH1 delet
981 30 88.2 20 2 AAW06885 Glycosami
982 30 88.2 20 2 AAR92270 Protein C
983 30 88.2 20 6 ABJ37568 Heparin b
984 30 88.2 20 9 ABB12129 Antimicro
985 30 88.2 51 4 AAU54014 Propionib
986 30 88.2 51 6 ABM50533 Propionib
987 30 88.2 89 4 ABB33444 Peptide #
988 30 88.2 89 4 ABG48291 Human liv
989 30 88.2 117 8 ADQ67108 Novel hum
990 30 88.2 120 4 AAB68965 Adenoviru
991 30 88.2 123 4 AAU19758 Human nov
992 30 88.2 123 4 AAU21495 Human nov
993 30 88.2 123 5 ABP47978 Human pol
994 30 88.2 123 7 ADC10940 Human ext
995 30 88.2 136 4 ABG12413 Novel hum
996 30 88.2 163 3 AAG33763 Arabidops
997 30 88.2 166 3 AAG33762 Arabidops
998 30 88.2 177 3 AAG33761 Arabidops
999 30 88.2 178 3 AAG54807 Arabidops
1000 30 88.2 181 6 ABM71724 Staphyloc

ALIGNMENTS

RESULT 1
AAB16978
ID AAB16978 standard; peptide; 6 AA.

AC AAB16978;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:34.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX autoimmune disease; cystostatic; antiasthmatic; thrombolytic; VEGF;
XX immunosuppressive; EPO; TPO; CRLA4; mimetic; IL-1; TNF; antagonist; MMP;
XX inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX thrombosis; pharmaceutical.

XX Synthetic.

OS WO200024782-A2.

PN 04-MAY-2000.

PD 25-OCT-1999; 99WO-US025044.

PF 23-OCT-1998; 98US-0105371P.

PR 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 207; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAG69443 to AAG69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. NO. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1 TLREWL 6

RESULT 2
ABB72864
ID ABB72864 standard; peptide; 6 AA.

AC ABB72864;

XX 05-APR-2002 (first entry)

DT TPO mimetic peptide SEQ ID NO:34.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
XX TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
XX TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
XX MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
XX cystostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
XX antianemic; anorectic; antiinfertility; haemostatic; dermatological;
XX neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
XX cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
XX sleep disorder; neurological degenerative disease; anaemia;
XX thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
XX Fanconi's syndrome.

OS Homo sapiens.

OS Synthetic.

PN WO200183525-A2.

PD 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

PA Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX diabetic retinopathy, obesity, sleep disorders and infertility.

PT Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX Sequence 6 AA;

XX Query Match 100.0%; Score 34; DB 5; Length 6;

XX Best Local Similarity 100.0%; Pred. No. 2e+06;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

DB 1 TLREWL 6

RESULT 3

ABG71746

ID ABG71746 standard; peptide; 6 AA.

XX AC ABG71746;

XX DT 20-JAN-2003 (first entry)

XX DE TPO receptor, MPL, agonist peptide consensus sequence #2.

XX Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;

XX complementarity determining region; CDR; antigenic; thrombopoietin; TPO;

XX thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;

XX T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;

XX proliferation; growth; differentiation; haematopoietic cell;

XX platelet progenitor cell; immune disorder; thrombocytopenia;

XX disseminated intravascular coagulation; stem cell; transplantation;

XX gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.

XX Synthetic.

XX WO200278612-A2.

XX PD 10-OCT-2002.

XX PF 02-APR-2002; 2002WO-US010301.

XX PR 02-APR-2001; 2001US-0281183P.

XX PA (PURD) PURDUE PHARMA LP.

XX PI Soltis DA, Burch RM, Ogert RA;

XX WPI; 2003-040615/03.

XX New thrombopoietin synthebodies, useful for stimulating proliferation,

PT growth, or differentiation of hematopoietic cells, for treating or

PT

PT preventing hematopoietic or immune disorders, e.g. thrombocytopenia.

XX Disclosure; Page 11; 97pp; English.

XX The invention discloses a variant of an immunoglobulin (Ig) variable

CC heavy or light chain domain that comprises at least one complementarity

CC determining region (CDR) and framework regions flanking the CDR. The CDR

CC also has added or substituted to it, at least one binding sequence which

CC is heterologous to the CDR and is an antigenic, agonistic sequence from a

CC thrombopoietin (TPO) receptor binding sequence. The antigenic sequence

CC can be a binding sequence heterologous to the CDR, a cytotoxic T-

CC lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper

CC cell sequence or a combination of each. The variant or thrombopoietin

CC synthebody, pharmaceutical and vaccine compositions are useful for

CC stimulating proliferation, growth or differentiation of haematopoietic

CC cells, particularly platelet progenitor cells. The variants are also

CC useful for treating or preventing haematopoietic or immune disorders

CC resulting from chemotherapy, radiation therapy, or bone marrow

CC transfusions (e.g. thrombocytopenia or disseminated intravascular

CC coagulation). Compositions comprising the synthebodies can be used for

CC the mobilisation, amplification and ex vivo expansion of stem cells and

CC committed precursor cells for autologous and allogeneic transplantation

CC as well as for the expansion of stem cells for gene therapy. They are

CC also useful as diagnostic or analytical reagents for studying the

CC function of thrombopoietin and its receptor in vivo or in vitro. The

CC sequence presented is the TPO receptor (MPL) agonist peptide consensus

CC sequence #2

XX Sequence 6 AA;

XX Query Match 100.0%; Score 34; DB 6; Length 6;

XX Best Local Similarity 100.0%; Pred. No. 2e+06;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

DB 1 TLREWL 6

RESULT 4

ADJ73015

ID ADJ73015 standard; peptide; 6 AA.

XX AC ADJ73015;

XX DT 06-MAY-2004 (first entry)

XX DE TPO mimetic peptide sequence SeqID 469.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;

XX cardiovascular; infectious; malignant; neurologic disease; anaemia;

XX immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

XX TPO.

XX Synthetic.

XX WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

PS Disclosure, SEQ ID NO 469; 97pp; English.

XX

XX This invention relates to novel mammalian CDR mimetibodies, specific portions or variants thereof. Specifically, it refers to an antibody fragment where a protein has been inserted into, or replaces a portion of, one or more CDR regions, such that each CDR mimetibody comprises at least one portion of a heavy chain or light chain variable region, which itself comprises at least one human framework region and at least one CDR region. The present invention describes human CDR mimetibodies, including modified immunoglobulins and cleavage products that can be useful in gene therapy and the generation of transgenic plants and animals. Furthermore, the CDR mimetibody is useful for preparing compositions for modulating, treating or reducing the symptoms of immune, cardiovascular, infectious, malignant and/or neurologic diseases, as well as anaemia. Accordingly, they exhibit immunomodulatory, cardiant, antimicrobial, cytostatic and neuroprotective activities. This peptide sequence is a TPO mimetic peptide sequence used to make a mimetibody of the invention.

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TLREWL 6
| | | | |
Db 1 TLREWL 6

RESULT 5
ADJ52650

ID ADJ52650 standard; peptide; 6 AA.

XX

AC ADJ52650;

XX

XX 06-MAY-2004 (first entry)

DT

DE CH1 deleted mimetibody-related peptide SeqID469.

XX

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant; hypotensive; neuroprotective; notropic; antibacterial; virucide; fungicide; gene therapy; immune disorder; cardiovascular disease; arrhythmia; hypertension; heart failure; neurodegenerative; multiple sclerosis; dementia; Alzheimer's disease; anaemia; cancerous condition; infectious disease; bacterial infection; viral infection; fungal infection.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002417-A2.

XX

PD 08-JAN-2004.

XX

XX 27-JUN-2003; 2003WO-US020347.

XX

XX 28-JUN-2002; 2002US-0392431P.

XX

XX (CENZ) CENTOCOR INC.

XX

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;
Kutoloski KA;

XX

XX WPI; 2004-082870/08.

XX

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for modulating, treating, alleviating, preventing an immune, cardiovascular, or neurodegenerative disease or disorder, anemia, cancer, or infectious diseases.

XX

PS Claim 2; SEQ ID NO 469; 129pp; English.

XX

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences which encode them), compositions, methods and uses. The invention may be useful for the development of compounds with an immunosuppressive, cardiovascular, cardiant, hypotensive, neuroprotective, notropic, antibacterial, virucide or fungicide activity. In addition, the disclosed sequences may prove useful for gene therapy. The CH1-deleted mimetibody is useful for diagnosing or treating a disease condition in a cell, tissue, organ or animal, specifically for modulating, treating, alleviating, preventing the incidence or reducing the symptoms of an immune, cardiovascular (for example arrhythmia, hypertension or heart failure), or neurodegenerative (for example multiple sclerosis, dementia or Alzheimer's disease) diseases or disorders, anaemia, cancerous conditions, or infectious diseases (for example bacterial, viral or fungal infection). The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.

XX

SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TLREWL 6
| | | | |
Db 1 TLREWL 6

RESULT 6
ADJ51611

ID ADJ51611 standard; peptide; 6 AA.

XX

AC ADJ51611;

XX

XX 06-MAY-2004 (first entry)

DT

DE CH1 deleted mimetibody-related peptide SeqID469.

XX

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen; dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen; gynaecological-Gen; hepatotropic; haemostatic; immunomodulator; antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic; antihemorrhagic; nephrotropic; respiratory-Gen; tumour necrosis factor; TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder; dental disorder; oral disorder; dermatological disorder; ear disorder; nose disorder; throat disorder; endocrine disorder; metabolic disorder; gastrointestinal disorder; gynaecological disorder; hepatic disorder; obstetric disorder; haematologic disorder; immunological disorder; allergic disorder; infectious disorder; musculoskeletal disorder; oncological disorder; neurological disorder; nutritional disorder; ophthalmologic disorder; pediatric disorder; psychiatric disorder; renal disorder; pulmonary disorder.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002424-A2.

XX

PD 08-JAN-2004.

XX

XX 30-JUN-2003; 2003WO-US020495.

XX

XX 28-JUN-2002; 2002US-0392431P.

PR

PR 19-SEP-2002; 2002US-0412144P.

XX

XX (CENZ) CENTOCOR INC.

XX

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;
Kutoloski KA;

XX

XX WPI; 2004-082872/08.

XX

PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic, and
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.
 XX
 PS Claim 14; SEQ ID NO 469; 123pp; English.
 XX
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 34; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
 DB 1 TLREWL 6

RESULT 7
 ADJ73040
 ID ADJ73040 standard; peptide; 9 AA.

XX
 AC ADJ73040;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE TPO mimetic peptide sequence SeqID 494.
 XX
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW TPO.

XX Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX
 XX New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 494; 97pp; English.
 PS
 CC This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is a TPO mimetic peptide sequence used to make a
 CC mimetibody of the invention.
 XX

Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
 DB 2 TLREWL 7

RESULT 8
 ADJ73037
 ID ADJ73037 standard; peptide; 9 AA.

XX AC ADJ73037;

XX 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 491.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW TPO.

XX Synthetic.

XX WO2003084477-A2.

XX 16-OCT-2003.

XX 24-MAR-2003; 2003WO-US009139.

XX 29-MAR-2002; 2002US-0368791P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 491; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7
|||||

RESULT 9

ID ADJ73041 standard; peptide; 9 AA.

XX AC ADJ73041;

DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 495.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX OS Synthetic.

XX PN WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX DR WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
XX cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 495; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7
|||||

RESULT 10

ID ADJ73039 standard; peptide; 9 AA.

XX AC ADJ73039;

XX DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 493.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX OS Synthetic.

XX PN WO2003084477-A2.

XX PD 16-OCT-2003.

XX PF 24-MAR-2003; 2003WO-US009139.

XX PR 29-MAR-2002; 2002US-0368791P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX DR WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
XX cardiovascular or neurologic disease.

XX PS Disclosure; SEQ ID NO 493; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 1 TLREWL 6
|||||
DB 2 TLREWL 7
|||||

RESULT 11
ADJ73038
ID ADJ73038 standard; peptide; 9 AA.
XX
AC ADJ73038;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 492.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
XX
WPI; 2003-804237/75.
XX
New CDR mimetibody comprising a portion of a heavy or light chain
variable region comprising human framework or ligand binding region,
useful for preparing a composition for treating e.g., immune,
cardiovascular or neurologic disease.
XX
Disclosure; SEQ ID NO 492; 97pp; English.
XX
This invention relates to novel mammalian CDR mimetibodies, specific
portions or variants thereof. Specifically, it refers to an antibody
fragment where a protein has been inserted into, or replaces a portion
of, one or more CDR regions, such that each CDR mimetibody comprises at
least one portion of a heavy chain or light chain variable region, which
itself comprises at least one human framework region and at least one
ligand binding region (LBR). The present invention describes human
CDR mimetibodies, including modified immunoglobulins and cleavage products
that can be useful in gene therapy and the generation of transgenic
plants and animals. Furthermore, the CDR mimetibody is useful for
preparing compositions for modulating, treating or reducing the symptoms
of immune, cardiovascular, infectious, malignant and/or neurologic
diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
cardiant, antimicrobial, cytostatic and neuroprotective activities. This
peptide sequence is a TPO mimetic peptide sequence used to make a
mimetibody of the invention.
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 34; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 1 TLREWL 6
|||||

DB 2 TLREWL 7
|||||

RESULT 12
AAW09471
ID AAW09471 standard; protein; 10 AA.
XX
AC AAW09471;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
PN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
PR 07-JUN-1995; 95US-00473604.
PR 07-JUN-1995; 95US-00476168.
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00484090.
PR 07-JUN-1995; 95US-00485301.
XX
(GLAX) GLAXO GROUP LTD.
XX
Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
WPI; 1997-051883/05.
XX
Thrombopoietin receptor-binding/activating peptide(s) and peptide
mimetic(s) - useful in treatment of haematological disorders, esp.
thrombocytopenia resulting from chemotherapy, etc.
XX
Disclosure; Page 26; 106pp; English.
XX
The present sequence is a peptide which binds to thrombopoietin (TPO)
receptor (TR). The compound can be used for treating patients suffering
from haematological disorders and thrombocytopenia resulting from
chemotherapy, radiation therapy or bone marrow transfusions. The peptide
may also be used to maintain the proliferation and growth of TPO-
dependent cell lines and for use in biological research, for detecting
TPO receptors on living cells
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
DB 5 TLREWL 10
|||||

RESULT 13
AAW09566
ID AAW09566 standard; protein; 10 AA.
XX
AC AAW09566;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX

KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX Synthetic.
XX OS
XX PN WO9640189-A1.
XX 19-DEC-1996.
XX PF
XX PF 05-JUN-1996; 96WO-US008998.
XX PR 07-JUN-1995; 95US-00472371.
XX PR 07-JUN-1995; 95US-00473604.
XX PR 07-JUN-1995; 95US-00476168.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00484090.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-051883/05.
XX XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 29; 106pp; English.
XX CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX SQ Sequence 10 AA;
XX
XX Query Match 100.0%; Score 34; DB 2; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 17;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 TLREWL 6
Db 2 TLREWL 7
XX
RESULT 14
AAW36622
ID AAW36622 standard; peptide; 10 AA.
XX AC AAW36622;
XX XX
XX 11-MAR-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX PN WO9640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 26; 106pp; English.
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX SQ Sequence 10 AA;
XX
XX Query Match 100.0%; Score 34; DB 2; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 17;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 TLREWL 6
Db 5 TLREWL 10
XX
RESULT 15
AAW36717
ID AAW36717 standard; peptide; 10 AA.
XX AC AAW36717;
XX XX
XX 27-FEB-1998 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX OS Synthetic.
XX PN WO9640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 29; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be used to treat disorders which are susceptible to treatment with a thrombopoietin agonist, preferably haematological disorders and thrombocytopenia resulting from chemotherapy, radiation therapy or bone marrow transfusions. It can also be used diagnostically, e.g. to investigate the mechanism of thrombopoietin signal transduction and receptor activation, or to maintain the proliferation and growth of thrombopoietin dependent cell lines

XX Sequence 10 AA;

Query Match 100.0%; Score 34; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7
|||||

RESULT 16
AAB17000
ID AAB17000 standard; peptide; 10 AA.
XX AC AAB17000;
XX 31-OCT-2000 (first entry)
XX TPO-mimetic peptide sequence SEQ ID NO:56.
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer; autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF; immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX vascular endothelial growth factor; matrix metalloproteinase; asthma; thrombosis; pharmaceutical.
XX Synthetic.
XX WO200024782-A2.
XX 04-MAY-2000.
XX 25-OCT-1999; 99WO-US025044.
XX 23-OCT-1998; 98US-0105371P.
XX 22-OCT-1999; 99US-00428082.
XX (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheatham J, Boone TC;
XX WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.
XX Claim 19; Page 214; 608pp; English.

XX The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAG69443 to AAG69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX Sequence 10 AA;

Query Match 100.0%; Score 34; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7
|||||

RESULT 17
AAU25841
ID AAU25841 standard; peptide; 10 AA.
XX AC AAU25841;
XX 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #27.
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine; haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA; bone marrow transplantation; haematological disorder; platelet disorder; enzyme-linked immunosorbent assay; in situ staining; biological fluid; tissue homogenate; fluorescence-activated cell sorting; Western blotting; in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX Homo sapiens.
XX US6251864-B1.
XX 26-JUN-2001.
XX 01-MAR-2000; 2000US-00516704.
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX 07-JUN-1996; 96WO-US009623.
XX 15-AUG-1996; 96US-00699027.
XX (GLAX) GLAXO GROUP LTD.
XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;
XX Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
XX Yin Q;
XX WPI; 2001-564142/63.
XX Activating thrombopoietin receptors in cells, used to treat thrombocytopenia and hematological disorders, comprises contacting cells with peptides and peptide mimetics attached to hydrophilic polymers.
XX Disclosure; Col 19; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 Db 5 TLREWL 10
 |||||

RESULT 18
 AAU25936
 ID AAU25936 standard; peptide; 10 AA.
 XX AC
 AC AAU25936;

DT 17-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #122.

DE Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;

XX WPI; 2001-564142/63.

DR Activating thrombopoietin receptors in cells, used to treat
 XX thrombocytopenia and hematological disorders, comprises contacting cells
 XX with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 22; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 Db 2 TLREWL 7
 |||||

RESULT 19
 ABB72886
 ID ABB72886 standard; peptide; 10 AA.
 XX AC
 AC ABB72886;

DT 05-APR-2002 (first entry)

XX TPO mimetic peptide SEQ ID NO:56.

DE Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IGG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TPO-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX Homo sapiens.

XX Synthetic.

XX WO200183525-A2.

XX 08-NOV-2001.

XX 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheatham JC, Boone TC, Gudas JM;
 PI WPI; 2002-130313/17.

XX Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 XX diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
 CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (II), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7

RESULT 20

ADJ52672
ID ADJ52672 standard; peptide; 10 AA.

XX AC ADJ52672;

XX DT 06-MAY-2004 (first entry)

XX DE CH1 deleted mimetibody-related peptide SeqID491.

XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX OS Unidentified.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1..10 /notes= "All Xaa's in this sequence are unidentified amino
FT acids"

XX FN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX XX 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner CA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX XX WPI; 2004-082870/08.

XX XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 491; 129pp; English.

XX PS This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 34; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7

RESULT 21

ADJ51633
ID ADJ51633 standard; peptide; 10 AA.

XX AC ADJ51633;

XX DT 06-MAY-2004 (first entry)

XX DE CH1 deleted mimetibody-related peptide SeqID491.

XX KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX OS Unidentified.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1..10 /notes= "All Xaa's in this sequence are unidentified amino
FT acids"

XX FN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX XX 28-JUN-2002; 2002US-0392431P.

XX XX 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;
PI Kutolowski KA;
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetide polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 14; SEQ ID NO 491; 123pp; English.
XX
XX This invention relates to CHI deleted mimetides (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, anti-allergic, muscular-Gen, cytostatic,
CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetide, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetide of the invention.
XX
XX Sequence 10 AA;
Query Match 100.0%; Score 34; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db 2 TLREWL 7
|||||
RESULT 22
AAB17001
ID AAB17001 standard; peptide; 11 AA.
XX
XX AAB17001;
AC
XX 31-OCT-2000 (first entry)
DT
XX TPO-mimetic peptide sequence SEQ ID NO:57.
DE
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
XX Synthetic.
OS
XX WO200024782-A2.
PN
XX 04-MAY-2000.
PD
XX 25-OCT-1999; 99WO-US025044.
PF
XX 23-OCT-1998; 98US-0105371P.
PR
XX 22-OCT-1999; 99US-00428082.
XX

PA (AMGE-) AMGEN INC.
XX
XX Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 19; Page 214; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAG59443 to AAG65526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
XX Sequence 11 AA;
Query Match 100.0%; Score 34; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
Db 2 TLREWL 7
|||||
RESULT 23
ABB72887
ID ABB72887 standard; peptide; 11 AA.
XX
XX ABB72887;
AC
XX 05-APR-2002 (first entry)
DT
XX TPO mimetic peptide SEQ ID NO:57.
DE
XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IGC; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX WO200183525-A2.
PN
XX 08-NOV-2001.
PD
XX 02-MAY-2001; 2001WO-US014310.
PF
XX

```

PR 03-MAY-2000; 2000US-00563286.
XX (ANGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX Claim 39; Page 43; 176pp; English.
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumor, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX Sequence 11 AA;
SQ
    Query Match      100.0%; Score 34; DB 5; Length 11;
    Best Local Similarity 100.0%; Pred. No. 19;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7

RESULT 24
ADJ52673
ID ADJ52673 standard; peptide; 11 AA.
XX
AC ADJ52673;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID492.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .11
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"

```

```

XX WO2004002417-A2.
XX 08-JAN-2004.
XX 27-JUN-2003; 2003WO-US020347.
XX 28-JUN-2002; 2002US-0392431P.
XX (CENZ ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghraryeb J, Scallon BJ, Neespor TC;
PI Kutolowski KA;
XX WPI; 2004-082870/08.
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX Claim 2; SEQ ID NO 492; 129pp; English.
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX Sequence 11 AA;
SQ
    Query Match      100.0%; Score 34; DB 8; Length 11;
    Best Local Similarity 100.0%; Pred. No. 19;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 2 TLREWL 7

RESULT 25
ADJ51634
ID ADJ51634 standard; peptide; 11 AA.
XX
AC ADJ51634;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID492.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulatory;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

```



```

XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 30
AAW36728
ID AAW36728 standard; peptide; 12 AA.
XX
AC AAW36728;
XX
XX 09-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 26; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 31
AAW36627
ID AAW36627 standard; peptide; 12 AA.
XX
AC AAW36627;
XX
XX 11-MAR-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 26; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 2 TLREWL 7

RESULT 32
AAW36675
ID AAW36675 standard; peptide; 12 AA.
XX
AC AAW36675;
XX
XX 27-FEB-1998 (first entry)
XX
XX Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX Synthetic.

```

```
XX WO9640750-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009623.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-052226/05.
XX PF Peptides and peptide mimetics which bind to and activate the
XX PT thrombopoietin receptor - useful in treatment of haematological
XX PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 28; 106pp; English.
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX CC used to treat disorders which are susceptible to treatment with a
XX CC thrombopoietin agonist, preferably haematological disorders and
XX CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX CC marrow transfusions. It can also be used diagnostically, e.g. to
XX CC investigate the mechanism of thrombopoietin signal transduction and
XX CC receptor activation, or to maintain the proliferation and growth of
XX CC thrombopoietin dependent cell lines
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB 2 TLREWL 7
RESULT 33
AAB17002
ID AAB17002 standard; peptide; 12 AA.
XX AC AAB17002;
XX DT 31-OCT-2000 (first entry)
XX DE TPO-mimetic peptide sequence SEQ ID NO:58.
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
XX KW immunosuppressive; EPO; TPO; C11A4; mimetic; IL-1; TNF; antagonist; MMP;
XX KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX KW thrombosis; pharmaceutical.
XX OS Synthetic.
XX PN WO200024782-A2.
XX PD 04-MAY-2000.
XX PF 25-OCT-1999; 99WO-US025044.
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham J, Boone TC;
XX PF Feige U, Liu C, Cheetham J, Boone TC;
XX DR WPI; 2000-350702/30.
XX PF Novel composition of matter comprising an Fc domain and pharmacologically
XX PT active peptides, useful for treating cancer and autoimmune diseases.
XX PS Claim 19; Page 214; 608pp; English.
XX CC The present invention describes composition of matter (I) comprising an
XX CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
XX CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
XX CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
XX CC P3, and P4 = are each independently sequences of pharmacologically active
XX CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
XX CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
XX CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
XX CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
XX CC cells from the present invention can be used for producing pharmaceutical
XX CC compositions. The compositions are useful for treating cancer, asthma,
XX CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
XX CC a Fab domain) can provide a longer half-life or incorporate functions
XX CC such as Fc receptor binding, protein A binding, complement fixation, and
XX CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
XX CC AAB18003 represent nucleotide and amino acid sequences used in the
XX CC exemplification of the present invention
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 34; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB 2 TLREWL 7
RESULT 34
AAB16993
ID AAB16993 standard; peptide; 12 AA.
XX AC AAB16993;
XX DT 31-OCT-2000 (first entry)
XX DE TPO-mimetic peptide sequence SEQ ID NO:49.
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
XX KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
XX KW immunosuppressive; EPO; TPO; C11A4; mimetic; IL-1; TNF; antagonist; MMP;
XX KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
XX KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
XX KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
XX KW thrombosis; pharmaceutical.
XX OS Synthetic.
XX PN WO200024782-A2.
XX PD 04-MAY-2000.
XX PF 25-OCT-1999; 99WO-US025044.
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham J, Boone TC;
```

```

DR WPI; 2000-350702/30.
XX
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
XX Claim 19; Page 212; 608pp; English.
XX
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
XX Sequence 12 AA;
SQ
Query Match 100.0%; Score 34; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db 2 TLREWL 7
|||||
|||||

RESULT 35
AAU25846
ID AAU25846 standard; peptide; 12 AA.
XX
XX AAU25846;
AC
XX
XX 17-DEC-2001 (first entry)
DT
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #32.
DE
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
XX Homo sapiens.
OS
XX
XX US6251864-B1.
PN
XX
XX 26-JUN-2001.
PD
XX
XX 01-MAR-2000; 2000US-00516704.
PF
XX
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00485301.
PR
XX 07-JUN-1996; 96WO-US009623.
PR
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
PA
XX
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX

```

```

DR WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 19; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 12 AA;
SQ
Query Match 100.0%; Score 34; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db 5 TLREWL 10
|||||
|||||

RESULT 36
AAU25951
ID AAU25951 standard; peptide; 12 AA.
XX
XX AAU25951;
AC
XX
XX 17-DEC-2001 (first entry)
DT
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #137.
DE
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
XX Homo sapiens.
OS
XX
XX US6251864-B1.
PN
XX
XX 26-JUN-2001.
PD
XX
XX 01-MAR-2000; 2000US-00516704.
PF
XX
XX 07-JUN-1995; 95US-00478128.
PR
XX 07-JUN-1995; 95US-00485301.
PR
XX 07-JUN-1996; 96WO-US009623.
PR
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
PA
XX
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX

```

XX WPI; 2001-564142/63.
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 XX Disclosure; Col 22; 128pp; English.
 XX
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 XX Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLRWL 6
 |||||
 Db 2 TLRWL 7

RESULT 37

AAU25894
 ID AAU25894 standard; peptide; 12 AA.

XX
 AC AAU25894;

XX
 DT 17-DEC-2001 (first entry)

XX
 DE Human thrombopoietin receptor (TPO-R) activator peptide #80.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; hematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-0048301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX

XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

PI Yin Q;

XX WPI; 2001-564142/63.
 XX

XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 21; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines

XX Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLRWL 6
 |||||
 Db 2 TLRWL 7

RESULT 38

AAU25947

ID AAU25947 standard; peptide; 12 AA.

XX
 AC AAU25947;

XX
 DT 17-DEC-2001 (first entry)

XX
 DE Human thrombopoietin receptor (TPO-R) activator peptide #133.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; hematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-0048301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX

XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 XX Disclosure; Col 22; 128pp; English.
 XX
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 34; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 |||||
 Db 2 TLREWL 7

RESULT 39
 ABB72888
 ID ABB72888 standard; peptide; 12 AA.
 XX
 AC ABB72888;

05-APR-2002 (first entry)

TPO mimetic peptide SEQ ID NO:58.

Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 cycostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 sleep disorder; neurological degenerative disease; anaemia;
 thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 Fanconi's syndrome.

Homo sapiens.
 Synthetic.

WO200183525-A2.

08-NOV-2001.

02-MAY-2001; 2001WO-US014310.

03-MAY-2000; 2000US-00563286.

(AMGE-) AMGEN INC.

Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

WPI; 2002-130313/17.

Novel vehicle-peptide molecule or its multimers useful for treating
 inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 diabetic retinopathy, obesity, sleep disorders and infertility.

Claim 39; Page 43; 176pp; English.

The present invention describes a vehicle-peptide molecule (I) or its
 multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 cycostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
 antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 neuroprotective activities. (I) can be used as a therapeutic or
 prophylactic agent as well as for screening purposes. (I) is useful for
 diagnosing diseases characterised by dysfunction of their associated
 protein of interest, for identifying normal or abnormal proteins of
 interest, as a part of diagnostic kit to detect the presence of their
 proteins of interest in a biological sample. Additionally, (I) is useful
 for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 infertility, and neurological degenerative diseases. (I), comprising EPO-
 mimetic compounds are useful for treating disorders characterised by low
 red blood cell levels such as anaemia. The TPO-mimetic comprising
 megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 tumour which result in thrombocytopenia, systemic lupus erythematosus,
 and Fanconi's syndrome. ABB72403 to ABB73426 and ABB35695 to ABB35777
 represent amino acid and nucleic acid sequences used in the
 exemplification of the present invention

Sequence 12 AA;

Query Match 100.0%; Score 34; DB 5; Length 12;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 |||||
 Db 2 TLREWL 7

RESULT 40

ABB72879
 ID ABB72879 standard; peptide; 12 AA.

XX ABB72879;

05-APR-2002 (first entry)

TPO mimetic peptide SEQ ID NO:49.

Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 cycostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 sleep disorder; neurological degenerative disease; anaemia;
 thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 Fanconi's syndrome.

Homo sapiens.
 Synthetic.


```

PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
XX Kutolowski KA;
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
XX or neurodegenerative disease or disorder, anemia, cancer, or infectious
XX diseases.
XX
XX Claim 2; SEQ ID NO 484; 129pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
XX which encode them), compositions, methods and uses. The invention may be
XX useful for the development of compounds with an immunosuppressive,
XX cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
XX antibacterial, virucide or fungicide activity. In addition, the disclosed
XX sequences may prove useful for gene therapy. The CHI-deleted mimetibody
XX is useful for diagnosing or treating a disease condition in a cell,
XX tissue, organ or animal, specifically for modulating, treating,
XX alleviating, preventing the incidence or reducing the symptoms of an
XX immune, cardiovascular (for example arrhythmia, hypertension or heart
XX failure), or neurodegenerative (for example multiple sclerosis, dementia
XX or Alzheimer's disease) diseases or disorders, anaemia, cancerous
XX conditions, or infectious diseases (for example bacterial, viral or
XX fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.
XX
XX Sequence 12 AA;
XX
XX Query Match 100.0%; Score 34; DB 8; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 20;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TLREWL 6
XX Db 2 TLREWL 7
XX
XX RESULT 43
XX ADJ52674
XX ID ADJ52674 standard; peptide; 12 AA.
XX
XX AC ADJ52674;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE CHI deleted mimetibody-related peptide SeqID493.
XX
XX KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
XX hypotensive; neuroprotective; nootropic; antibacterial; virucide;
XX fungicide; gene therapy; immune disorder; cardiovascular disease;
XX arrhythmia; hypertension; heart failure; neurodegenerative;
XX multiple sclerosis; dementia; Alzheimer's disease; anaemia;
XX cancerous condition; infectious disease; bacterial infection;
XX viral infection; fungal infection.
XX
XX OS Unidentified.
XX OS Synthetic.
XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 1..12
XX FT /note= "All Xaa's in this sequence are unidentified amino
XX FT acids"
XX
XX PN WO2004002417-A2.
XX

```

```

PD 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
XX Kutolowski KA;
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
XX modulating, treating, alleviating, preventing an immune, cardiovascular,
XX or neurodegenerative disease or disorder, anemia, cancer, or infectious
XX diseases.
XX
XX Claim 2; SEQ ID NO 493; 129pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
XX which encode them), compositions, methods and uses. The invention may be
XX useful for the development of compounds with an immunosuppressive,
XX cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
XX antibacterial, virucide or fungicide activity. In addition, the disclosed
XX sequences may prove useful for gene therapy. The CHI-deleted mimetibody
XX is useful for diagnosing or treating a disease condition in a cell,
XX tissue, organ or animal, specifically for modulating, treating,
XX alleviating, preventing the incidence or reducing the symptoms of an
XX immune, cardiovascular (for example arrhythmia, hypertension or heart
XX failure), or neurodegenerative (for example multiple sclerosis, dementia
XX or Alzheimer's disease) diseases or disorders, anaemia, cancerous
XX conditions, or infectious diseases (for example bacterial, viral or
XX fungal infection). The present sequence is that of a peptide which may be
XX used during the creation of a mimetibody of the invention.
XX
XX Sequence 12 AA;
XX
XX Query Match 100.0%; Score 34; DB 8; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 20;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TLREWL 6
XX Db 2 TLREWL 7
XX
XX RESULT 44
XX ADJ51626
XX ID ADJ51626 standard; peptide; 12 AA.
XX
XX AC ADJ51626;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE CHI deleted mimetibody-related peptide SeqID484.
XX
XX KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
XX gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
XX anti-allergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
XX ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
XX TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
XX dental disorder; oral disorder; dermatological disorder; ear disorder;
XX nose disorder; throat disorder; endocrine disorder; metabolic disorder;
XX gastrointestinal disorder; gynaecological disorder; hepatic disorder;
XX obstetric disorder; haematologic disorder; immunological disorder;
XX allergic disorder; infectious disorder; musculoskeletal disorder;
XX oncological disorder; neurological disorder; nutritional disorder;
XX ophthalmologic disorder; pediatric disorder; psychiatric disorder;
XX renal disorder; pulmonary disorder.
XX
XX OS Unidentified.
XX

```


OS Synthetic.

XX WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for

PT diagnosing, preventing or treating cardiovascular, dermatologic,

PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 14; SEQ ID NO 484; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CHI deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, haematologic, gastrointestinal, gynaecological, hepatic,

CC obstetric, metabolic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 2 TLREWL 7

RESULT 45

ADJ51635

ID ADJ51635 standard; peptide; 12 AA.

XX ADJ51635;

XX 06-MAY-2004 (first entry)

XX CHI deleted mimetibody-related peptide SeqID493.

XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;

XX dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

XX gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

XX antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

XX ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

XX TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

XX dental disorder; oral disorder; dermatological disorder; ear disorder;

XX nose disorder; throat disorder; endocrine disorder; metabolic disorder;

XX gastrointestinal disorder; gynaecological disorder; hepatic disorder;

XX obstetric disorder; haematologic disorder; immunological disorder;

XX allergic disorder; infectious disorder; musculoskeletal disorder;

XX oncological disorder; neurological disorder; nutritional disorder;

XX ophthalmologic disorder; pediatric disorder; psychiatric disorder;

XX renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1.12

FT /note= "All Xaa's in this sequence are unidentified amino

FT acids"

XX WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

PI Kutoloski KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for

PT diagnosing, preventing or treating cardiovascular, dermatologic,

PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 14; SEQ ID NO 493; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulator, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CHI deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, haematologic, gastrointestinal, gynaecological, hepatic,

CC obstetric, metabolic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 34; DB 8; Length 12;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

Db 2 TLREWL 7

RESULT 46

AAW36732

ID AAW36732 standard; peptide; 13 AA.

PN WO200024782-A2.
 XX 04-MAY-2000.
 XX 25-OCT-1999; 99WO-US025044.
 XX 23-OCT-1998; 98US-0105371P.
 XX 22-OCT-1999; 99US-00428082.
 XX (AMGE-) AMGEN INC.
 XX Feige U, Liu C, Cheetham J, Boone TC;
 XX WPI; 2000-350702/30.
 XX Novel composition of matter comprising an Fc domain and pharmacologically
 XX active peptides, useful for treating cancer and autoimmune diseases.
 XX Claim 19; Page 215; 608pp; English.
 XX The present invention describes composition of matter (I) comprising an
 XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 XX (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 XX (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 XX P3, and P4 = are each independently sequences of pharmacologically active
 XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 XX c, d, e, and f = are each independently 0 or 1, provided that at least 1
 XX of a and b is 1. The composition can have cytostatic, antiasthmatic,
 XX thrombolytic and immunosuppressive activities. DNAs, vectors and host
 XX cells from the present invention can be used for producing pharmaceutical
 XX compositions. The compositions are useful for treating cancer, asthma,
 XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 XX a Fab domain) can provide a longer half-life or incorporate functions
 XX such as Fc receptor binding, protein A binding, complement fixation, and
 XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 XX AAB18003 represent nucleotide and amino acid sequences used in the
 XX exemplification of the present invention
 XX Sequence 13 AA;
 XX Query Match 100.0%; Score 34; DB 3; Length 13;
 XX Best Local Similarity 100.0%; Pred. No. 22;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 Db 2 TLREWL 7
 RESULT 49
 ABB72901
 ID ABB72901 standard; peptide; 13 AA.
 XX ABB72901;
 XX 05-APR-2002 (first entry)
 XX TPO mimetic peptide SEQ ID NO:71.
 XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 XX TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 XX TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 XX MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 XX cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;
 XX antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 XX neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 XX cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 XX sleep disorder; neurological degenerative disease; anaemia;
 XX thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 XX Fanconi's syndrome.
 OS Homo sapiens.
 OS Synthetic.
 PN WO200183525-A2.
 XX 08-NOV-2001.
 XX 02-MAY-2001; 2001WO-US014310.
 XX 03-MAY-2000; 2000US-00563286.
 XX (AMGE-) AMGEN INC.
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX WPI; 2002-130313/17.
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 XX diabetic retinopathy, obesity, sleep disorders and infertility.
 XX Claim 39; Page 44; 176pp; English.
 XX The present invention describes a vehicle-peptide molecule (I) or its
 XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 XX cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
 XX antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 XX neuroprotective activities. (I) can be used as a therapeutic or
 XX prophylactic agent as well as for screening purposes. (I) is useful for
 XX diagnosing diseases characterised by dysfunction of their associated
 XX protein of interest, for identifying normal or abnormal proteins of
 XX interest, as a part of diagnostic kit to detect the presence of their
 XX proteins of interest in a biological sample. Additionally, (I) is useful
 XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 XX infertility, and neurological degenerative diseases. (I), comprising EPO-
 XX mimetic compounds are useful for treating disorders characterised by low
 XX red blood cell levels such as anaemia. The TPO-mimetic comprising
 XX compounds are useful for treating conditions that involve an existing
 XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 XX deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 XX tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 XX and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 XX represent amino acid and nucleic acid sequences used in the
 XX exemplification of the present invention
 XX Sequence 13 AA;
 XX Query Match 100.0%; Score 34; DB 5; Length 13;
 XX Best Local Similarity 100.0%; Pred. No. 22;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 Db 6 TLREWL 11
 RESULT 50
 ABB72889
 ID ABB72889 standard; peptide; 13 AA.
 XX ABB72889;
 XX 05-APR-2002 (first entry)
 XX TPO mimetic peptide SEQ ID NO:59.
 XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 XX erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 XX TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 XX TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 XX MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 XX cytostatic; antirheumatic; antiarthritis; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX WO200183525-A2.
 XX
 XX
 PD 08-NOV-2001.
 XX
 XX
 PF 02-MAY-2001; 2001WO-US014310.
 XX
 XX
 PR 03-MAY-2000; 2000US-00563286.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 FI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX
 DR WPI; 2002-130313/17.
 XX
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 XX
 XX Claim 39; Page 43; 176pp; English.

The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The EPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

SQ Sequence 13 AA;

Query Match 100.0%; Score 34; DB 5; Length 13;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 |||||
 Db 2 TLREWL 7

Search completed: May 12, 2006, 10:36:42
 Job time : 49.6897 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 7.28205 Seconds

(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-34

Perfect score: 34

Sequence: 1 TLREWL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR_80:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	486	2 B86411	protein F3M18.4 [i
2	34	100.0	1045	2 S23570	pol polyprotein ho
3	32	94.1	473	2 E84853	hypothetical prote
4	32	94.1	560	2 S41808	glucose-6-phosphat
5	32	94.1	566	2 T09154	glucose-6-phosphat
6	32	94.1	568	2 S57830	glucose-6-phosphat
7	32	94.1	569	2 S23542	glucose-6-phosphat
8	32	94.1	569	2 S41806	glucose-6-phosphat
9	32	94.1	569	2 S57831	pol polyprotein ho
10	32	94.1	1295	2 S60179	hypothetical prote
11	31	91.2	85	2 F75291	hypothetical prote
12	31	91.2	105	2 T38761	conserved hypotet
13	31	91.2	158	2 A10605	hypothetical prote
14	31	91.2	158	2 E64823	probable sensory t
15	31	91.2	158	2 E90745	hypothetical prote
16	31	91.2	287	2 C83872	conserved hypotet
17	31	91.2	299	2 B82535	nitrogenase cofact
18	31	91.2	384	2 S34816	transforming growt
19	31	91.2	409	2 S01825	transforming growt
20	31	91.2	410	2 A41397	transforming growt
21	31	91.2	410	2 A55706	transforming growt
22	31	91.2	412	2 A34939	transforming growt
23	31	91.2	412	2 A36169	transforming growt
24	31	91.2	415	1 O4P8CP	campbor 5-monooxyg
25	31	91.2	505	2 T19971	hypothetical prote
26	31	91.2	506	2 T19973	hypothetical prote
27	31	91.2	554	2 A96018	probable regulator
28	31	91.2	639	2 A56126	peroxisomal target
29	31	91.2	936	2 E96508	hypothetical prote

30	31	91.2	984	2 T48216	hypothetical prote
31	31	91.2	1094	2 F70897	probable arabinosy
32	31	91.2	1630	2 T00390	KIAA614 protein -
33	31	91.2	4342	2 H83343	probable non-ribo
34	30	88.2	120	2 S10865	early E4 13K prote
35	30	88.2	158	2 D72305	hypothetical prote
36	30	88.2	181	2 G89799	hypothetical prote
37	30	88.2	233	2 S52461	hypothetical prote
38	30	88.2	233	2 S38228	hypothetical prote
39	30	88.2	233	2 S53104	hypothetical prote
40	30	88.2	239	2 E69189	hypothetical prote
41	30	88.2	263	2 A30227	hypothetical prote
42	30	88.2	284	2 F83042	hypothetical prote
43	30	88.2	292	2 F86431	hypothetical prote
44	30	88.2	308	2 S77938	RNA-LP protein -
45	30	88.2	319	2 F82601	5'-nucleotidase XF
46	30	88.2	332	2 A29711	deacetoxycephalosp
47	30	88.2	338	2 E90998	hypothetical prote
48	30	88.2	346	2 D85818	unknown protein en
49	30	88.2	394	2 Ar2963	phosphomannose iso
50	30	88.2	394	2 G98319	mannose-6-phosphat
51	30	88.2	406	2 A39339	protein C inhibito
52	30	88.2	427	2 B83001	probable C4-dicarb
53	30	88.2	497	2 T15812	hypothetical prote
54	30	88.2	515	2 G70941	hypothetical prote
55	30	88.2	521	2 T11166	CDPdiacylglycerol-
56	30	88.2	521	2 T01923	hypothetical prote
57	30	88.2	658	1 A65079	arginine decarboxy
58	30	88.2	658	2 AG0876	arginine decarboxy
59	30	88.2	658	2 F91105	biosynthetic argin
60	30	88.2	658	2 A85951	biosynthetic argin
61	30	88.2	659	2 A80114	arginine decarboxy
62	30	88.2	952	2 E86147	arginine decarboxy
63	30	88.2	1040	2 T08190	hypothetical prote
64	30	88.2	1094	2 C70612	hypothetical prote
65	30	88.2	1123	2 T28139	probable recB prot
66	30	88.2	6260	2 T30228	PK4 protein kinase
67	30	88.2	8563	2 T30226	polyketide synthas
68	29	85.3	65	2 S59524	polyketide synthas
69	29	85.3	80	2 H82194	triose-phosphate i
70	29	85.3	84	2 T03339	conserved hypotet
71	29	85.3	106	1 Z28PT9	gene e28 protein -
72	29	85.3	114	2 A10294	gene 49.2 protein
73	29	85.3	124	2 AG3149	probable membrane
74	29	85.3	124	2 B98138	hypothetical prote
75	29	85.3	143	2 E71041	hypothetical prote
76	29	85.3	165	2 A70326	conserved hypotet
77	29	85.3	192	2 B81911	probable molybdot
78	29	85.3	199	2 T03101	hypothetical prote
79	29	85.3	201	2 G59096	hypothetical prote
80	29	85.3	232	2 B82987	probable hydrolase
81	29	85.3	243	2 S38236	hypothetical prote
82	29	85.3	251	2 T35258	probable transcrip
83	29	85.3	255	2 F64503	hypothetical prote
84	29	85.3	269	2 S71999	hypothetical prote
85	29	85.3	277	2 G71456	hypothetical prote
86	29	85.3	283	2 A64174	hypothetical prote
87	29	85.3	283	2 T11147	hypothetical prote
88	29	85.3	285	2 H72073	phospho-2-dehydro-
89	29	85.3	285	2 H86550	deoxyheptonate ald
90	29	85.3	289	2 E83454	cell division prot
91	29	85.3	301	2 C82394	probable glycine c
92	29	85.3	308	2 S58504	reverse transcript
93	29	85.3	327	2 D82090	conserved hypotet
94	29	85.3	337	2 F91191	probable LPS biosy
95	29	85.3	337	2 G86038	probable LPS biosy
96	29	85.3	338	2 AD1142	tagatose-1,6-dipho
97	29	85.3	338	2 AG1500	tagatose-1,6-dipho
98	29	85.3	339	2 C86978	hypothetical prote
99	29	85.3	339	2 F70901	riboflavin bifunct
100	29	85.3	341	2 E65080	hypothetical prote
101	29	85.3	341	2 F91107	probable protein t
102	29	85.3	341	2 E85952	probable protein t

103 29 85.3 344 2 E84377 protein export [im
104 29 85.3 350 2 AG0879 A/G-specific adeni
105 29 85.3 350 2 A40647 methylation-indepe
106 29 85.3 353 2 A10036 conserved hypotet
107 29 85.3 357 2 E83205 carotenoid biosynt
108 29 85.3 359 1 F64407 probable iron-sulf
109 29 85.3 364 1 C69351 hypotetrical prote
110 29 85.3 369 2 C75043 membrane-bound lyt
111 29 85.3 370 2 A82578 outer membrane lip
112 29 85.3 370 2 A97360 peptide methionine
113 29 85.3 394 2 C82439 hypotetrical prote
114 29 85.3 399 2 D96832 hypotetrical prote
115 29 85.3 414 2 H97601 aminotransferase I
116 29 85.3 414 2 A12823 N-ethylmaline ch
117 29 85.3 416 2 E69374 RNA polymerase sig
118 29 85.3 417 1 S67566 probable membrane
119 29 85.3 423 2 E86132 hypotetrical prote
120 29 85.3 423 2 C91291 hypotetrical prote
121 29 85.3 426 2 S6561 ATP synthase in ty
122 29 85.3 440 2 A83435 hypotetrical prote
123 29 85.3 441 2 C64026 virion morphogenes
124 29 85.3 454 2 F64124 light-harvesting I
125 29 85.3 482 2 T50901 probable amidaae p
126 29 85.3 485 2 E83620 adult-specific 6l.
127 29 85.3 540 2 B45655 hypotetrical prote
128 29 85.3 541 2 AF2483 probable glucose-6
129 29 85.3 567 2 T03948 glucose-6-phosphat
130 29 85.3 567 2 S02094 hypotetrical prote
131 29 85.3 570 2 S09812 hypotetrical prote
132 29 85.3 576 2 H70961 CIP synthase (EC 6
133 29 85.3 597 2 T00471 dopamine transport
134 29 85.3 619 2 S20346 dopamine transport
135 29 85.3 619 2 T59558 dopamine transport
136 29 85.3 620 2 I59737 dopamine transport
137 29 85.3 620 2 A48980 NADH2 dehydrogenas
138 29 85.3 633 2 B70946 probable r8bU prot
139 29 85.3 653 2 C70742 3',5'-cyclic-nucle
140 29 85.3 713 2 JW0088 topoisomerase iv c
141 29 85.3 750 2 A97501 topoisomerase IV c
142 29 85.3 750 2 A82719 topoisomerase IV s
143 29 85.3 775 2 A83402 hypotetrical prote
144 29 85.3 816 2 A71006 mannosyltransferas
145 29 85.3 860 2 C82750 hypotetrical prote
146 29 85.3 897 2 T21688 alanine-tRNA ligas
147 29 85.3 924 2 E71476 hypotetrical prote
148 29 85.3 1003 2 T34066 neuronal different
149 29 85.3 1019 2 JC7538 DNA-directed DNA p
150 29 85.3 1092 2 H69071 probable membrane
151 29 85.3 1150 2 S49956 hypotetrical prote
152 29 85.3 1204 2 T18812 probable type II D
153 29 85.3 1252 2 D71810 type IIS restricti
154 29 85.3 1279 2 B64709 tegument protein h
155 29 85.3 1369 2 T03104 hypotetrical prote
156 29 85.3 1545 2 T26589 6-methylsalicylic
157 29 85.3 1774 2 S13178 pyrimidine synthe
158 29 85.3 2225 1 A23443 hypotetrical prote
159 29 85.3 55 2 T30476 hypotetrical prote
160 29 85.3 105 2 PW0018 hypotetrical prote
161 29 85.3 105 2 D72546 hypotetrical prote
162 29 85.3 120 1 QXR93 nonstructural prot
163 29 85.3 140 2 AD1205 molybdopterin conv
164 29 85.3 140 2 AC1562 molybdopterin conv
165 29 85.3 147 2 H82172 heat shock protein
166 29 85.3 182 2 A40227 CDPdiacylglycerol-
167 29 85.3 209 1 XURTWC methylated-DNA-lpr
168 29 85.3 244 2 A82316 hypotetrical prote
169 29 85.3 251 2 A83390 arginine/orntine
170 29 85.3 258 2 T08265 conserved hypotet
171 29 85.3 267 2 T36675 probable integral
172 29 85.3 279 2 H81678 phospho-2-dehydro-
173 29 85.3 317 2 A72073 hypotetrical prote
174 29 85.3 317 2 F86552 hypotetrical prote

176 28 82.4 176 protein export [im
177 28 82.4 177 A/G-specific adeni
178 28 82.4 178 methylation-indepe
179 28 82.4 179 conserved hypotet
180 28 82.4 180 carotenoid biosynt
181 28 82.4 181 probable iron-sulf
182 28 82.4 182 hypotetrical prote
183 28 82.4 183 membrane-bound lyt
184 28 82.4 184 outer membrane lip
185 28 82.4 185 peptide methionine
186 28 82.4 186 hypotetrical prote
187 28 82.4 187 hypotetrical prote
188 28 82.4 188 aminotransferase I
189 28 82.4 189 N-ethylmaline ch
190 28 82.4 190 RNA polymerase sig
191 28 82.4 192 1 S67566 probable membrane
192 28 82.4 193 hypotetrical prote
193 28 82.4 194 hypotetrical prote
194 28 82.4 195 hypotetrical 48.2K
195 28 82.4 196 ATP synthase in ty
196 28 82.4 197 hypotetrical prote
197 28 82.4 198 virion morphogenes
198 28 82.4 199 light-harvesting I
199 28 82.4 200 probable amidaae p
200 28 82.4 201 adult-specific 6l.
201 28 82.4 202 hypotetrical prote
202 28 82.4 203 probable glucose-6
203 28 82.4 204 glucose-6-phosphat
204 28 82.4 205 hypotetrical prote
205 28 82.4 206 hypotetrical prote
206 28 82.4 207 CIP synthase (EC 6
207 28 82.4 208 dopamine transport
208 28 82.4 209 dopamine transport
209 28 82.4 210 dopamine transport
210 28 82.4 211 NADH2 dehydrogenas
211 28 82.4 212 probable r8bU prot
212 28 82.4 213 3',5'-cyclic-nucle
213 28 82.4 215 topoisomerase iv c
214 28 82.4 216 topoisomerase IV s
215 28 82.4 217 topoisomerase IV c
216 28 82.4 217 hypotetrical prote
217 27 79.4 218 mannolyltransferas
218 27 79.4 219 hypotetrical prote
219 27 79.4 220 hypotetrical prote
220 27 79.4 221 alanine-tRNA ligas
221 27 79.4 222 hypotetrical prote
222 27 79.4 223 neuronal different
223 27 79.4 224 DNA-directed DNA p
224 27 79.4 225 probable membrane
225 27 79.4 226 hypotetrical prote
226 27 79.4 227 probable type II D
227 27 79.4 228 type IIS restricti
228 27 79.4 229 tegument protein h
229 27 79.4 230 hypotetrical prote
230 27 79.4 231 6-methylsalicylic
231 27 79.4 232 pyrimidine synthe
232 27 79.4 233 hypotetrical prote
233 27 79.4 234 hypotetrical prote
234 27 79.4 235 hypotetrical prote
235 27 79.4 236 nonstructural prot
236 27 79.4 237 molybdopterin conv
237 27 79.4 238 molybdopterin conv
238 27 79.4 239 heat shock protein
239 27 79.4 240 CDPdiacylglycerol-
240 27 79.4 241 methylated-DNA-lpr
241 27 79.4 242 hypotetrical prote
242 27 79.4 243 arginine/orntine
243 27 79.4 244 conserved hypotet
244 27 79.4 245 probable integral
245 27 79.4 246 phospho-2-dehydro-
246 27 79.4 247 hypotetrical prote
247 27 79.4 248 hypotetrical prote

probable transcrip
hypotetrical prote
conserved hypotet
hypotetrical prote
activin - fruit fl
probable membrane-
hypotetrical membr
hypotetrical prote
probable RND efflu
protein F2809.3 [1
folypolyglutamate
inner membrane pro
hypotetrical prote
hypotetrical prote
glucose-6-phosphat
hypotetrical prote
homeotic protein B
transposase alr801
DNA topoisomerase
two component sens
probable bacteriop
methylmalonyl-CoA
conserved hypotet
conserved hypotet
hypotetrical prote
hypotetrical prote
sensory transducti
pyruvate, water di
hypotetrical prote
cell cycle histidi
conserved hypotet
replication licens
probable polA prot
hypotetrical prote
glutamyl aminopept
hypotetrical prote
protein F2624.8 [1
carbamoyl-phosphat
probable molybdopt
excisionase [impor
probable integrase
hypotetrical prote
hypotetrical prote
orotate phosphorib
hypotetrical prote
hypotetrical prote
conserved hypotet
Bola protein [limpo
early E4 13K prote
60s ribosomal prot
hypotetrical prote
transcription regu
ribosomal protein
hypotetrical prote
secretin precursor
hypotetrical prote
hypotetrical prote
hypotetrical prote
hypotetrical prote
hypotetrical prote
alpha-lactalbumin
hypotetrical prote
hypotetrical prote
molybdopterin bios
probable sensory t
beta-fructofuranos
dihydrofolate redu
hypotetrical prote
hypotetrical prote
CDPdiacylglycerol-

249	27	79.4	182	2	AC0749	CPDdiacylglycerol-	322	27	79.4	321	1	BVECBF	biotin-lacetyl-CoA
250	27	79.4	182	2	B90960	CPDdiacylglycerol-	323	27	79.4	321	2	A86089	hypothetical prote
251	27	79.4	182	2	C85808	CPDdiacylglycerol-	324	27	79.4	321	2	D91241	hypothetical prote
252	27	79.4	182	2	AG2062	hypothetical prote	325	27	79.4	324	2	C82251	GGDEF family prote
253	27	79.4	185	2	F64049	CDP-diacylglycerol	326	27	79.4	325	1	G90393	phosphoesterase-re
254	27	79.4	185	2	H82227	CDP-diacylglycerol	327	27	79.4	325	2	T04562	catechol O-methyl
255	27	79.4	186	2	B83322	CDP-diacylglycerol	328	27	79.4	329	2	T19872	hypothetical prote
256	27	79.4	187	2	C81803	CDP-diacylglycerol	329	27	79.4	330	2	F82338	hypothetical prote
257	27	79.4	187	2	F81085	CDP-diacylglycerol	330	27	79.4	335	2	D71275	hypothetical prote
258	27	79.4	187	2	A82746	conserved hypothet	331	27	79.4	336	2	A64378	hypothetical prote
259	27	79.4	189	2	A12653	molybdenum cofacto	332	27	79.4	336	2	T26189	hypothetical prote
260	27	79.4	189	2	E97435	molybdopterin bios	333	27	79.4	337	2	E69136	hypothetical prote
261	27	79.4	189	2	A83586	probable transcrip	334	27	79.4	337	2	T28690	hypothetical prote
262	27	79.4	196	2	S40728	hypothetical prote	335	27	79.4	345	2	D88485	protein F23F12.4 [
263	27	79.4	198	2	S73275	hypothetical prote	336	27	79.4	350	2	B38535	A/G-specific adeni
264	27	79.4	206	1	I40173	orotate phosphorib	337	27	79.4	350	2	H85953	adenine glycosylas
265	27	79.4	206	2	AD1898	urease accessory p	338	27	79.4	350	2	E91108	photosystem II pro
266	27	79.4	207	2	G97633	thymidine kinase (339	27	79.4	361	1	F65012	hypothetical prote
267	27	79.4	209	2	AG1303	orotate phosphorib	340	27	79.4	361	2	A91037	probable peptidase
268	27	79.4	209	2	AG1675	CDP-diacylglycerol	342	27	79.4	361	2	C85881	probable peptidase
269	27	79.4	211	2	G82573	Nicotinate-nucleot	343	27	79.4	361	2	T25196	hypothetical prote
270	27	79.4	213	1	E64798	hypothetical prote	344	27	79.4	363	2	E90268	acyl-CoA dehydroge
271	27	79.4	213	2	A85564	hypothetical prote	345	27	79.4	365	2	T25309	hypothetical prote
272	27	79.4	213	2	E90713	hypothetical prote	346	27	79.4	367	2	T43252	carbamoyl-phosphat
273	27	79.4	215	2	T18842	hypothetical prote	347	27	79.4	369	2	T40905	WD repeat protein
274	27	79.4	216	2	H69953	conserved hypothet	348	27	79.4	376	2	T51170	homocitrate synth
275	27	79.4	216	2	T18176	hypothetical prote	349	27	79.4	376	2	T35868	probable dipeptida
276	27	79.4	221	2	A82190	hypothetical prote	350	27	79.4	376	2	A75449	phospho-2-dehydro-
277	27	79.4	226	2	H82710	phosphoglycolate p	351	27	79.4	377	2	C70628	hypothetical prote
278	27	79.4	227	2	T49725	hypothetical prote	352	27	79.4	378	2	F90323	coenzyme PQQ synth
279	27	79.4	230	2	H86758	hypothetical prote	353	27	79.4	378	2	D90347	formate dehydrogen
280	27	79.4	231	2	E81215	thiol-disulfide in	354	27	79.4	379	2	J02272	hypothetical prote
281	27	79.4	234	2	T46203	transcription fact	355	27	79.4	379	2	F71533	catechol O-methyl
282	27	79.4	241	2	F95355	protein [imported	356	27	79.4	382	2	T04963	gamma-butyrobetain
283	27	79.4	243	2	B72577	probable phosphate	357	27	79.4	383	2	S32457	conserved hypothet
284	27	79.4	244	2	T07068	beta-fructofuranos	358	27	79.4	384	2	E81689	trehalose-phosphat
285	27	79.4	246	2	B55445	int protein - Esch	359	27	79.4	385	1	T52057	corrinoid/iron-aui
286	27	79.4	246	2	T00231	recombinase - Esch	360	27	79.4	389	2	B69096	hypothetical prote
287	27	79.4	247	2	S18604	triiose-phosphate i	361	27	79.4	391	2	A71936	conserved hypothet
288	27	79.4	248	2	A81040	conserved hypothet	362	27	79.4	392	2	E72219	conserved hypothet
289	27	79.4	248	2	S56531	hypothetical prote	363	27	79.4	392	2	B64648	pyruvate synthase
290	27	79.4	248	2	A12175	hypothetical prote	364	27	79.4	396	2	D69174	3-hydroxy-3-methyl
291	27	79.4	253	1	ISZMT	triiose-phosphate i	365	27	79.4	397	2	D45088	hypothetical prote
292	27	79.4	260	2	G83989	hydrolase BH2719 [366	27	79.4	404	2	B64543	probable transposa
293	27	79.4	261	2	T32399	hypothetical prote	367	27	79.4	409	2	D75629	glucose-1-phosphat
294	27	79.4	262	2	C75554	conserved hypothet	368	27	79.4	413	2	JV0087	periplasmic glucos
295	27	79.4	262	2	S55487	probable 3-methyl-	369	27	79.4	413	2	F90773	periplasmic glucos
296	27	79.4	267	2	A98141	hypothetical prote	370	27	79.4	414	2	C75461	hypothetical prote
297	27	79.4	270	2	A31416	glucose-1-phosphat	371	27	79.4	416	2	D75407	probable transposa
298	27	79.4	275	2	B81430	probable prephenat	372	27	79.4	417	2	S07286	hypothetical prote
299	27	79.4	283	2	A86923	probable glycosyl	373	27	79.4	420	2	B87201	hypothetical prote
300	27	79.4	288	2	S73016	H+-transporting tw	374	27	79.4	420	2	A70796	probable cycloprop
301	27	79.4	290	2	B97100	pyridoxal kinase r	375	27	79.4	420	2	E75378	probable valine-py
302	27	79.4	292	2	F84951	hypothetical prote	376	27	79.4	423	2	A55797	collagen precursor
303	27	79.4	293	2	T00466	coatomer complex e	377	27	79.4	423	2	S63583	tetrahydrofolylpol
304	27	79.4	294	2	T10858	probable carboxype	378	27	79.4	424	2	B81936	tetrahydrofolylpol
305	27	79.4	295	2	AB3613	glucose-1-phosphat	379	27	79.4	424	2	C81170	folylpolyglutamate
306	27	79.4	296	2	AG0147	probable membrane	380	27	79.4	431	2	JW0098	carbazole dioxygen
307	27	79.4	297	1	JN0773	calponin H1 - rat	381	27	79.4	431	2	C75354	conserved hypothet
308	27	79.4	297	1	S31484	calponin H1 - pig	382	27	79.4	433	2	D69270	ferredoxin-nitrite
309	27	79.4	297	1	S31486	calponin H1 - mous	383	27	79.4	436	2	A38145	invariant surface
310	27	79.4	297	2	JC4500	basic calponin - h	384	27	79.4	436	2	C82682	glutamate-cysteine
311	27	79.4	297	2	G02142	smooth muscle cell	385	27	79.4	454	2	AG1221	ethanolamine ammon
312	27	79.4	297	2	S55063	conserved hypothet	386	27	79.4	454	2	AB1575	ethanolamine ammon
313	27	79.4	299	2	A75401	ribulose-phosphate	387	27	79.4	454	2	A97048	similar to phospho
314	27	79.4	301	2	S62416	n-acetylglucosamin	388	27	79.4	464	2	H33006	hypothetical prote
315	27	79.4	302	2	T36030	probable fructokin	389	27	79.4	466	2	JC6525	processing peptida
316	27	79.4	309	2	S23096	protein kinase [EC	390	27	79.4	468	2	H70521	probable papA2 pro
317	27	79.4	310	2	B55171	chloroplast import	391	27	79.4	468	2	AD1926	hypothetical prote
318	27	79.4	315	2	T18696	hypothetical prote	392	27	79.4	469	2	AD1926	hypothetical prote
319	27	79.4	315	2	B98226	hypothetical prote	393	27	79.4	470	2	AB2188	phage-related prot
320	27	79.4	315	2	AF3060	conserved hypothet	394	27	79.4	470	2	D82575	
321	27	79.4	320	2	AG0934	bifunctional prote							

395	27	79.4	471	2	B38637	Ras inhibitor (clo	468	27	79.4	1919	2	T42098	PAM interacting pr
396	27	79.4	472	2	B82545	phage-related prot	469	27	79.4	2137	2	T05244	hypothetical prote
397	27	79.4	473	2	T06167	beta-fructofuranos	470	27	79.4	2157	2	AI3009	peptide synthetase
398	27	79.4	481	2	T35149	two-component syst	471	27	79.4	2186	2	TL1169	tiggrin - fruit fl
399	27	79.4	485	2	S56359	probable proton/ol	472	27	79.4	2233	1	TLN2P3	genome polyprotein
400	27	79.4	485	2	F86108	probable peptide t	473	27	79.4	2255	1	JQ1532	genome polyprotein
401	27	79.4	485	2	H91267	probable peptide t	474	27	79.4	2298	2	T49648	hypothetical prote
402	27	79.4	499	2	A02068	cell death suppres	475	27	79.4	2326	2	T29140	hypothetical prote
403	27	79.4	505	2	B83918	carboxypeptidase B	476	27	79.4	2566	2	E98274	hypothetical prote
404	27	79.4	509	2	T30861	traJ protein - Sal	477	27	79.4	2761	2	T29285	hypothetical prote
405	27	79.4	513	1	S50216	translation initia	478	27	79.4	3163	2	AB0233	yersinibactin bio
406	27	79.4	540	2	T76869	hypothetical prote	479	27	79.4	3163	2	T17440	probable polyketid
407	27	79.4	542	2	AB1066	conserved hypotet	480	27	79.4	4136	2	T43274	dynamin heavy chain
408	27	79.4	553	2	B83640	hypothetical prote	481	27	79.4	15281	2	S41309	cyclosporin synthet
409	27	79.4	555	2	F71083	hypothetical prote	482	26	76.5	35	2	H64642	hypothetical prote
410	27	79.4	555	2	A43784	Xpo protein - Afri	483	26	76.5	70	2	T06920	ribosomal protein
411	27	79.4	562	2	F72771	probable lysyl-tRN	484	26	76.5	72	1	RSBEXL	excisionase [phag
412	27	79.4	575	2	C86398	protein T7N9.26 [i	485	26	76.5	72	1	A90729	excisionase [impor
413	27	79.4	581	2	T12095	beta-fructofuranos	486	26	76.5	72	2	S06533	excisionase - phag
414	27	79.4	581	2	F84599	probable kinesin h	487	26	76.5	72	2	C70241	hypothetical prote
415	27	79.4	584	2	S51882	topoisomerase I-re	488	26	76.5	73	2	H71893	probable molybdopt
416	27	79.4	590	2	D69722	thiamin biosynthes	489	26	76.5	74	2	A64620	molybdopterin conv
417	27	79.4	592	2	S56681	beta-fructofuranos	490	26	76.5	82	1	CCP85D	cytochrome c551 -
418	27	79.4	592	2	S25705	Ig mu chain - shee	491	26	76.5	87	2	S42171	cytochrome-c oxida
419	27	79.4	595	2	B83891	thiamin biosynthes	492	26	76.5	87	2	T10462	hypothetical prote
420	27	79.4	595	2	T39228	beta-transducin -	493	26	76.5	96	2	H87336	conserved hypotet
421	27	79.4	606	1	UZAD12	terminal protein p	494	26	76.5	101	2	F72573	hypothetical prote
422	27	79.4	611	2	F75095	probable asparagin	495	26	76.5	103	2	T07053	cysteine proteinas
423	27	79.4	618	2	B87564	conserved hypotet	496	26	76.5	106	2	JQ0234	hypothetical 12.5K
424	27	79.4	623	2	T48859	disease resistance	497	26	76.5	107	2	S50516	hypothetical prote
425	27	79.4	625	2	S60401	nuclear division p	498	26	76.5	108	1	RPECW	trp operon repress
426	27	79.4	634	1	GERTX1	matrix glycoprotei	499	26	76.5	108	1	S21331	hypothetical prote
427	27	79.4	655	2	T26061	hypothetical prote	500	26	76.5	108	2	S45254	trp operon repress
428	27	79.4	664	2	S60062	hevin precursor -	501	26	76.5	108	2	G91297	regulator for trp
429	27	79.4	670	2	S76872	hypothetical prote	502	26	76.5	108	2	A86139	trp operon repress
430	27	79.4	676	1	A39379	hatching-suppresse	503	26	76.5	109	2	S45257	trp operon repress
431	27	79.4	686	2	F89895	ATP-dependent DNA	504	26	76.5	121	2	A70555	hypothetical prote
432	27	79.4	693	2	AE0005	ATP-dependent DNA	505	26	76.5	123	1	LACM	alpha-lactalbumin
433	27	79.4	716	1	BVECAD	membrane protein t	506	26	76.5	125	2	JU0179	heparin-binding pr
434	27	79.4	738	2	S10659	ABC transporter, A	507	26	76.5	125	2	AC0610	probable integrase
435	27	79.4	757	2	B75437	hypothetical prote	508	26	76.5	127	2	B86158	60S ribosomal prot
436	27	79.4	763	2	S96571	hypothetical prote	509	26	76.5	128	1	S52084	ribosomal protein
437	27	79.4	788	1	B4AG58	virB4 protein prec	510	26	76.5	128	2	JC2120	heparin-binding pr
438	27	79.4	789	1	B4AGN6	component of type	511	26	76.5	128	2	JC2121	heparin-binding pr
439	27	79.4	789	1	AB3248	phenylalanine-tRNA	512	26	76.5	128	2	JC2119	heparin-binding pr
440	27	79.4	795	2	H84944	acetyl-CoA decarbo	513	26	76.5	130	2	A30033	development-specif
441	27	79.4	798	2	F69549	phosphoenolpyruvat	514	26	76.5	130	2	I83571	probable membrane
442	27	79.4	803	2	B82392	calpain (EC 3.4.22	515	26	76.5	130	2	C85701	hypothetical prote
443	27	79.4	805	1	A55054	calpain (EC 3.4.22	516	26	76.5	130	2	F90843	hypothetical prote
444	27	79.4	805	2	A39764	carbon-monoxide de	517	26	76.5	131	2	S74539	hypothetical prote
445	27	79.4	809	2	B83409	hypothetical prote	518	26	76.5	133	2	S45752	hypothetical prote
446	27	79.4	822	2	AD3232	conjugal transfer	519	26	76.5	137	2	A87586	hypothetical prote
447	27	79.4	834	2	T06055	hypothetical prote	520	26	76.5	137	2	B96603	transcription fact
448	27	79.4	848	2	C81356	aconitate hydratase	521	26	76.5	138	2	C84068	polyribonucleotide
449	27	79.4	850	2	C83081	probable oxidoredu	522	26	76.5	138	2	T40987	hypothetical prote
450	27	79.4	852	2	F44020	hypothetical prote	523	26	76.5	142	2	A82966	conserved hypotet
451	27	79.4	852	2	T08502	trbE protein - Ent	524	26	76.5	143	2	A70953	hypothetical prote
452	27	79.4	906	2	G96621	probable disease r	525	26	76.5	143	2	C75173	hypothetical prote
453	27	79.4	968	2	T47546	hypothetical prote	526	26	76.5	144	2	F97571	hypothetical prote
454	27	79.4	978	2	T14968	phage lambda-relat	527	26	76.5	144	2	AG2792	conserved hypotet
455	27	79.4	1009	2	S44621	C50C3.2 protein -	528	26	76.5	148	2	S26450	hypothetical prote
456	27	79.4	1014	2	JE0333	klotko protein - r	529	26	76.5	148	2	AD3139	hypothetical prote
457	27	79.4	1047	2	S19508	MSH3 protein - yea	530	26	76.5	154	2	E75024	n-cell receptor be
458	27	79.4	1056	2	T28636	zinc-metallopeptid	531	26	76.5	156	2	C48230	hypothetical prote
459	27	79.4	1173	2	H83189	DNA polymerase III	532	26	76.5	156	2	H84251	flagella-related p
460	27	79.4	1175	2	T46124	hypothetical prote	533	26	76.5	157	2	B69476	inosine monophosph
461	27	79.4	1221	2	T18550	reverse gyrase cha	534	26	76.5	158	2	S43892	DNA-directed DNA p
462	27	79.4	1272	2	C90593	hypothetical prote	535	26	76.5	158	2	G64749	hypothetical prote
463	27	79.4	1349	2	T13031	nucleoporin - frui	536	26	76.5	161	2	A69732	PASX prophage ORF
464	27	79.4	1365	2	T13991	nucleoporin 154 -	537	26	76.5	162	2	G70005	hypothetical prote
465	27	79.4	1541	1	S71839	canaliculac multid	538	26	76.5	163	2	T36985	hypothetical prote
466	27	79.4	1608	1	WMTQM	183K protein - tob	539	26	76.5	164	2	AF3320	conserved hypotet
467	27	79.4	1899	2	T32732	PAM C-terminal int	540	26	76.5	166	2	H64416	VPS29-like phospho

541	26	76.5	166	2	T40988	probable glutaredo	614	26	76.5	253	2	C84644	probable ribonucle
542	26	76.5	174	2	G83712	hypothetical prote	615	26	76.5	254	1	TRWV3Y	trypsin-like prote
543	26	76.5	175	2	D85544	primosomal replica	616	26	76.5	257	2	C70715	hypothetical prote
544	26	76.5	175	2	H30693	primosomal replica	617	26	76.5	260	2	S17300	cytochrome-c oxida
545	26	76.5	175	2	JQ1149	primosomal replica	618	26	76.5	260	2	S14157	cytochrome-c oxida
546	26	76.5	176	2	B64549	conserved hypothe	619	26	76.5	260	2	A82837	conserved hypothe
547	26	76.5	176	2	F83531	hypothetical prote	620	26	76.5	260	2	A12480	hypothetical prote
548	26	76.5	177	2	A82294	hypothetical prote	621	26	76.5	262	2	C91111	hypothetical prote
549	26	76.5	177	2	H87412	hypothetical prote	622	26	76.5	263	2	F84694	endo-1,3-1,4-beta-
550	26	76.5	178	2	E91080	hypothetical prote	623	26	76.5	263	2	A83055	trypsin-like prote
551	26	76.5	178	2	F85925	hypothetical prote	624	26	76.5	263	2	A83055	hypothetical prote
552	26	76.5	180	2	G70028	conserved hypothe	625	26	76.5	267	2	A32794	probable tropinone
553	26	76.5	183	2	AG1003	probable oxidoredu	626	26	76.5	267	2	AC3133	probable tropinone
554	26	76.5	183	2	AI0023	conserved hypothe	627	26	76.5	268	2	E84695	hypothetical prote
555	26	76.5	184	1	B65129	probable NAD(P)H2	628	26	76.5	269	2	C84314	hypothetical prote
556	26	76.5	184	2	G85999	probable NAD(P)H o	629	26	76.5	269	2	H84378	hypothetical prote
557	26	76.5	184	2	B91154	probable NAD(P)H o	630	26	76.5	269	2	S58439	transcription fact
558	26	76.5	184	2	H72414	hypothetical prote	631	26	76.5	274	2	E83444	hypothetical prote
559	26	76.5	184	2	H84143	hypothetical prote	632	26	76.5	274	2	T26658	hypothetical prote
560	26	76.5	185	2	E70029	conserved hypothe	633	26	76.5	276	2	B70598	PBSX prophage prot
561	26	76.5	185	2	D82838	conserved hypothe	634	26	76.5	277	2	I40410	DNA-directed DNA p
562	26	76.5	186	2	G91158	hypothetical prote	635	26	76.5	278	2	S06416	probable deoxyhept
563	26	76.5	186	2	E86004	hypothetical prote	636	26	76.5	278	2	B71522	hypothetical prote
564	26	76.5	186	2	H65134	hypothetical 21.2	637	26	76.5	279	2	T26166	hypothetical prote
565	26	76.5	187	2	C82065	hypothetical prote	638	26	76.5	279	2	T21068	hypothetical prote
566	26	76.5	188	2	AD0533	conserved hypothe	639	26	76.5	279	2	T15662	hypothetical prote
567	26	76.5	189	2	B69020	conserved hypothe	640	26	76.5	281	2	A43749	lignin beta-ether
568	26	76.5	190	2	A82343	yaed protein - Eec	641	26	76.5	281	2	F85956	probable transposa
569	26	76.5	191	1	H64744	hypothetical prote	642	26	76.5	286	2	G70537	hypothetical prote
570	26	76.5	191	2	F84932	hypothetical prote	643	26	76.5	289	2	C88231	endo-1,3-1,4-beta-
571	26	76.5	191	2	B85505	probable phosphata	644	26	76.5	290	2	A35519	isochorismatase (S
572	26	76.5	191	2	B90654	conserved hypothe	645	26	76.5	291	1	JH0489	phosphoribosylamin
573	26	76.5	192	2	G95072	conserved hypothe	646	26	76.5	291	2	S55291	phosphoribosylamin
574	26	76.5	192	2	D97940	conserved hypothe	647	26	76.5	291	2	S55292	phosphoribosylamin
575	26	76.5	193	2	A86852	hypothetical prote	648	26	76.5	293	2	S55293	phosphoribosylamin
576	26	76.5	193	2	B84382	hypothetical prote	649	26	76.5	294	2	S22613	xylose operon regu
577	26	76.5	198	2	AD0999	probable NUDIX hyd	650	26	76.5	294	2	AC1420	hypothetical prote
578	26	76.5	199	2	E85635	part of trimethyla	651	26	76.5	294	2	AD1795	hypothetical prote
579	26	76.5	199	2	A90773	part of trimethyla	652	26	76.5	294	2	A96155	probable transcrip
580	26	76.5	200	2	T23485	hypothetical prote	653	26	76.5	295	2	H91117	probable transcrip
581	26	76.5	202	2	AC2875	hypothetical prote	654	26	76.5	295	2	G85962	phosphoribosylamin
582	26	76.5	203	2	S57836	lymphoid-specific	655	26	76.5	301	2	JC2039	phosphoribosylamin
583	26	76.5	204	2	E90668	probable membrane	656	26	76.5	306	1	QJ1395	conserved hypothe
584	26	76.5	204	2	H85518	probable membrane	657	26	76.5	307	2	AB0562	conserved hypothe
585	26	76.5	204	2	G64754	probable membrane	658	26	76.5	308	2	AC1867	hypothetical prote
586	26	76.5	205	2	T16560	hypothetical prote	659	26	76.5	311	2	G64238	hypothetical prote
587	26	76.5	207	2	T37464	probable glutathio	660	26	76.5	311	2	AG3526	2-hydroxyhepta-2,4
588	26	76.5	208	2	AD2808	conserved hypothe	661	26	76.5	316	2	D83406	probable transmemb
589	26	76.5	209	2	B87627	hypothetical prote	662	26	76.5	316	2	T32962	hypothetical prote
590	26	76.5	211	2	B75470	hydrolase, CbhY/Cb	663	26	76.5	317	2	E95239	conserved hypothe
591	26	76.5	213	2	AC3389	uracil-DNA glycosy	664	26	76.5	317	2	E98103	probable transcrip
592	26	76.5	215	2	E96533	hypothetical prote	665	26	76.5	317	2	G83544	probable transcrip
593	26	76.5	217	2	B97587	hypothetical prote	666	26	76.5	321	2	S86866	probable transcrip
594	26	76.5	217	2	T35633	hypothetical prote	667	26	76.5	322	2	H87131	probable lytic mur
595	26	76.5	218	2	B71183	hypothetical prote	668	26	76.5	323	2	G82135	pectic enzyme secr
596	26	76.5	218	2	D72073	conserved hypothe	669	26	76.5	324	2	D37874	ribosomal large ch
597	26	76.5	218	2	D86550	CT383 hypotheical	670	26	76.5	324	2	G82289	hypothetical prote
598	26	76.5	218	2	A82765	hypothetical prote	671	26	76.5	326	2	S73828	probable transposa
599	26	76.5	219	2	E75143	phosphoglycolate p	672	26	76.5	327	2	E82277	probable transposa
600	26	76.5	223	2	A86249	protein T23J18.24	673	26	76.5	328	1	DWSMGG	replication protei
601	26	76.5	225	2	D83200	conserved hypothe	674	26	76.5	328	2	A35257	hypothetical prote
602	26	76.5	228	2	C87467	ABC transporter, A	675	26	76.5	328	2	C70745	hypothetical prote
603	26	76.5	228	2	A13547	carboxylesterase (676	26	76.5	331	2	B90173	transcription acti
604	26	76.5	229	2	S64049	hypothetical prote	677	26	76.5	331	2	S45584	probable transcrip
605	26	76.5	236	2	D72355	conserved hypothe	678	26	76.5	332	2	AD2541	transcription inic
606	26	76.5	236	2	AF3416	glutathione transf	679	26	76.5	332	2	E75332	probable ADP-ribos
607	26	76.5	237	2	JC7392	5'TG3'-interacting	680	26	76.5	333	2	E97257	spore coat protein
608	26	76.5	246	2	B27271	hypothetical prote	681	26	76.5	333	2	I59348	CCAAT binding tran
609	26	76.5	248	1	QJ1682	infected cell proc	682	26	76.5	334	2	T51106	hypothetical prote
610	26	76.5	249	2	H86786	hypothetical prote	683	26	76.5	335	2	T21461	hypothetical prote
611	26	76.5	250	2	T14548	beta-fructofuranos	684	26	76.5	336	2	T27635	homeobox protein c
612	26	76.5	253	2	S53760	triiose-phosphate i	685	26	76.5	337	2	T27635	conserved hypothe
613	26	76.5	253	2	E72302	esterase - Thermot	686	26	76.5	339	2	G87213	conserved hypothe

833	26	76.5	624	2	S74952	hypothetical prote	906	26	76.5	853	2	AD0544	outer membrane fim
834	26	76.5	625	2	T40742	hypothetical integ	907	26	76.5	858	1	VCLJG2	env polyprotein pr
835	26	76.5	626	2	T15972	hypothetical prote	908	26	76.5	859	1	VCLJCT	DNA-directed RNA p
836	26	76.5	653	1	UZAD22	terminal protein p	909	26	76.5	867	2	G69485	leucine-trNA ligas
837	26	76.5	657	1	UZADP5	terminal protein p	910	26	76.5	869	2	S77455	leucyl-trNA synthe
838	26	76.5	657	1	A64079	2',3'-cyclic-nucle	911	26	76.5	872	2	AD2216	leucyl-trNA synthe
839	26	76.5	662	2	T24906	sucrose 1P-fructos	912	26	76.5	874	2	D75307	hypothetical prote
840	26	76.5	664	2	T24459	hypothetical prote	913	26	76.5	886	2	T24169	cation-transportin
841	26	76.5	670	2	T02092	beta-fructofuranos	914	26	76.5	902	2	C83964	hypothetical prote
842	26	76.5	671	2	S32230	Ca2+-transporting	915	26	76.5	906	2	T22376	hypothetical prote
843	26	76.5	681	1	H82059	2',3'-cyclic-nucle	916	26	76.5	908	2	D87434	phosphoenolpyruvat
844	26	76.5	682	1	JH0560	cyclic nucleotide-	917	26	76.5	909	2	A43802	cellulase (EC 3.2.
845	26	76.5	688	2	T44343	hypothetical prote	918	26	76.5	927	1	JQ0948	AS antigen precurs
846	26	76.5	691	2	A54741	erythrocyte membra	919	26	76.5	951	2	T08987	hypothetical prote
847	26	76.5	693	2	A12584	hypothetical prote	920	26	76.5	954	2	H71427	hypothetical prote
848	26	76.5	697	2	H85205	potassium channel	921	26	76.5	955	2	T00247	zinc finger protei
849	26	76.5	697	2	T04931	potassium channel	922	26	76.5	955	2	A47055	transposase - Erwi
850	26	76.5	698	1	JX0202	long-chain-fatty-a	923	26	76.5	961	2	T15680	hypothetical prote
851	26	76.5	702	2	C82792	methionyl-trNA syn	924	26	76.5	967	2	T15680	hypothetical prote
852	26	76.5	718	2	C82817	ATP-dependent DNA	925	26	76.5	968	2	C82452	probable leus prot
853	26	76.5	721	2	A39707	erythrocyte membra	926	26	76.5	969	2	A70912	leucine-trNA ligas
854	26	76.5	728	2	H97366	ribonucleotide red	927	26	76.5	972	2	T10023	probable leucyl tr
855	26	76.5	739	2	B82552	phage-related prot	928	26	76.5	973	2	A85055	hypothetical prote
856	26	76.5	739	2	F82769	phage-related tail	929	26	76.5	980	2	T27342	hypothetical prote
857	26	76.5	742	2	A82357	hypothetical prote	930	26	76.5	981	2	T41029	probable aminoacid
858	26	76.5	743	2	B84639	probable ubiquitin	931	26	76.5	985	1	VCLJSF	env polyprotein -
859	26	76.5	744	2	S57061	nuclear pore prote	932	26	76.5	985	1	VCLJSF	zinc proteinase (E
860	26	76.5	761	2	G70393	conserved hypotet	933	26	76.5	989	2	T46183	probable sensor/re
861	26	76.5	762	2	H83415	cis/trans isomeraz	934	26	76.5	992	2	A83324	leucine-trNA ligas
862	26	76.5	764	2	A47456	down-regulated in	935	26	76.5	994	1	SYNCLM	DNA-directed DNA p
863	26	76.5	766	2	H83141	probable two-compo	936	26	76.5	994	2	S18739	env protein - siml
864	26	76.5	769	1	ISBYT1	DNA topoisomerase	937	26	76.5	995	2	C84145	hypothetical prote
865	26	76.5	772	2	E82592	transcription-rela	938	26	76.5	999	2	AG2413	hypothetical prote
866	26	76.5	773	2	T40694	probable rna matur	939	26	76.5	1003	2	S07526	Ca2+-transporting
867	26	76.5	786	2	H69980	single-strand DNA-	940	26	76.5	1003	2	TI3856	ksr protein - frui
868	26	76.5	791	2	T28815	hypothetical prote	941	26	76.5	1018	2	AG0703	conserved hypotet
869	26	76.5	792	2	A00104	probable glucosida	942	26	76.5	1018	2	B90928	probable oxidase l
870	26	76.5	793	2	T40285	hypothetical prote	943	26	76.5	1018	2	F85776	probable oxidase y
871	26	76.5	795	1	SVECFB	phenylalanine-trNA	944	26	76.5	1018	2	AH0293	conserved hypotet
872	26	76.5	795	2	AE0705	phenylalanine-trNA	945	26	76.5	1018	2	AG0293	conserved iron-sulf
873	26	76.5	795	2	H85779	phenylalanine trNA	946	26	76.5	1021	2	F82230	conserved hypotet
874	26	76.5	795	2	D90931	phenylalanine trNA	947	26	76.5	1027	2	B64187	conserved hypotet
875	26	76.5	795	2	D82225	phenylalanine-trNA	948	26	76.5	1029	2	T00712	protein kinase hom
876	26	76.5	795	2	AD0296	phenylalanine-trNA	949	26	76.5	1039	2	S02711	cellulase (EC 3.2.
877	26	76.5	803	2	AD1282	leucyl-trNA synthe	950	26	76.5	1039	2	S02711	1-phosphatidylinos
878	26	76.5	803	2	AH1653	leucyl-trNA synthe	951	26	76.5	1068	1	A43322	1-phosphatidylinos
879	26	76.5	804	2	B89961	leucyl-trNA synthe	952	26	76.5	1068	1	I38110	kinesin-like prote
880	26	76.5	804	2	D69650	leucine-trNA ligas	953	26	76.5	1075	2	T45570	hypothetical prote
881	26	76.5	805	2	T48840	transducer protein	954	26	76.5	1075	2	D70568	hypothetical prote
882	26	76.5	806	2	A84060	leucyl-trNA synthe	955	26	76.5	1089	2	E81446	hypothetical prote
883	26	76.5	807	2	F85647	probable outer mem	956	26	76.5	1100	2	T17270	hsp70-related prot
884	26	76.5	807	2	F64844	ycds protein precu	957	26	76.5	1112	2	T02848	dynein heavy chain
885	26	76.5	807	2	F90787	probable outer mem	958	26	76.5	1114	2	T30299	protein-tyrosine-p
886	26	76.5	810	2	C84845	probable salt-indu	959	26	76.5	1139	1	A49724	hypothetical prote
887	26	76.5	812	2	D96979	leucyl-trNA synthe	960	26	76.5	1139	2	B70954	parasporal crystal
888	26	76.5	815	2	T36671	probable helicase	961	26	76.5	1172	2	S32689	transcription-repa
889	26	76.5	819	2	G81698	leucyl-trNA synthe	962	26	76.5	1179	2	AG1101	transcription-repa
890	26	76.5	819	2	C71544	leucine-trNA ligas	963	26	76.5	1179	2	AG1463	protein F15H18.21
891	26	76.5	820	2	A86510	leucyl trNA synthe	964	26	76.5	1199	2	D86317	probable pol polyP
892	26	76.5	820	2	C72113	leucine-trNA ligas	965	26	76.5	1220	1	T18348	DNA-directed DNA p
893	26	76.5	822	2	E75523	ATP-dependent heli	966	26	76.5	1220	2	DJBEC3	DNA-directed DNA p
894	26	76.5	824	2	T10020	leucine-trNA ligas	967	26	76.5	1220	2	T42573	DNA-directed DNA p
895	26	76.5	824	2	F72408	leucine-trNA ligas	968	26	76.5	1235	1	DJBEC3	DNA-directed DNA p
896	26	76.5	825	2	JC4163	DNA-binding protel	969	26	76.5	1235	1	DJBEC7	DNA-directed DNA p
897	26	76.5	829	2	H86726	leucine-trNA ligas	970	26	76.5	1235	1	DJBEC1	DNA-directed DNA p
898	26	76.5	833	2	H95029	leucyl-trNA synthe	971	26	76.5	1240	1	DJBEC1	microcystin synthe
899	26	76.5	833	2	C97901	leucine-trNA ligas	972	26	76.5	1271	2	DJBEC21	hypothetical prote
900	26	76.5	837	2	C69200	surface proteinase	973	26	76.5	1313	2	F96673	mannan endo-1,4-be
901	26	76.5	840	2	B69216	probable proteinase	974	26	76.5	1331	2	A48954	DNA-directed RNA p
902	26	76.5	844	2	C91212	probable outer mem	975	26	76.5	1391	2	F81233	DNA-directed RNA p
903	26	76.5	844	2	C86058	probable fibmbrial	976	26	76.5	1391	2	F81233	DNA-directed RNA p
904	26	76.5	849	2	C87740	protein H26D21.2 [977	26	76.5	1391	2	C82007	DNA-directed RNA p
905	26	76.5	851	2	AE3567	ATP-dependent heli	978	26	76.5	1391	2	C82007	DNA-directed RNA p

979 26 76.5 1438 2 T17402
 980 26 76.5 1453 2 F88640
 981 26 76.5 1476 1 A39901
 982 26 76.5 1476 1 A40303
 983 26 76.5 1561 2 T00248
 984 26 76.5 1683 2 AF2071
 985 26 76.5 1691 1 A44212
 986 26 76.5 1711 2 T31337
 987 26 76.5 1711 2 AD1842
 988 26 76.5 1742 2 T17120
 989 26 76.5 1747 2 AC1842
 990 26 76.5 1755 2 F82618
 991 26 76.5 1808 2 T47792
 992 26 76.5 1858 1 A44214
 993 26 76.5 2262 2 S16664
 994 26 76.5 2327 2 T31733
 995 26 76.5 2731 1 VF1HJH
 996 26 76.5 2733 2 S15760
 997 26 76.5 3161 2 T30342
 998 26 76.5 4063 2 T42993
 999 26 76.5 4101 2 T23630
 1000 26 76.5 5825 2 T12117

ALIGNMENTS

RESULT 1
 B86411
 protein F3M18.4 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: B86411
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
 aneen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: B86411
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-486 <STO>
 A;Cross-references: UNIPROT:Q9SGQ0; UNIPARC:UPI000000ABA54; GB:AB005172; NID:g6560771; PI
 A;Gene: F3M18.4
 A;Map position: 1

Query Match 100.0%; Score 34; DB 2; Length 486;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
 |||||
 Db 311 TLREWL 316

RESULT 2
 S23570
 pol polyprotein homolog - fungus (Cladosporium fulvum) retrotransposon Cft-1 (fragment)
 N;Alternate names: reverse transcriptase
 C;Species: Cladosporium fulvum
 C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000
 C;Accession: S23570; S19850
 R;McHale, M.T.; Roberts, I.N.; Noble, S.M.; Beaumont, C.; Whitehead, M.P.; Seth, D.; Oli
 Mol. Gen. Genet. 233, 337-347, 1992
 A;Title: Cft-1: an LTR-retrotransposon in Cladosporium fulvum, a fungal pathogen of toma
 A;Reference number: S23569; MUID:92318885; PMID:1377773

A;Accession: S23570
 A;Molecule type: DNA
 A;Residues: 1-1045 <MCH>
 A;Cross-references: UNIPARC:UPI000011E65B; EMBL:Z11866; NID:g2562; PIDN:CAA77891.1; PID:
 C;Genetics:
 A;Mobile element: retrotransposon Cft-1
 C;Keywords: polyprotein

Query Match 100.0%; Score 34; DB 2; Length 1045;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
 |||||
 Db 336 TLREWL 341

RESULT 3
 E84853
 hypothetical protein At2g42400 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C;Accession: E84853
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: E84853
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-473 <STO>
 A;Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI000017A02D; GB:AB002093; NID:g4567312; PI
 C;Genetics:
 A;Gene: At2g42400
 A;Map position: 2

Query Match 94.1%; Score 32; DB 2; Length 473;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
 |||||
 Db 347 TLREWL 352

RESULT 4
 S41808
 Glucose-6-phosphate isomerase (EC 5.3.1.9), cytosolic - Arabidopsis thaliana
 N;Alternate names: phosphoglucose isomerase
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C;Accession: S41808
 R;Thomas, B.R.; Ford, V.S.; Pichersky, E.; Gottlieb, L.D.
 Genetics 135, 895-905, 1993
 A;Title: Molecular characterization of duplicate cytosolic phosphoglucose isomerase gene
 A;Reference number: S41806; MUID:94123980; PMID:8293986
 A;Accession: S41808
 A;Molecule type: DNA
 A;Residues: 1-560 <THO>
 A;Cross-references: UNIPARC:UPI0000001752; EMBL:X69195; NID:g415922; PIDN:CAA48940.1; PI
 C;Genetics:
 A;Gene: pgic
 A;Introns: 18/3; 44/2; 71/3; 87/3; 139/3; 172/1; 186/2; 209/1; 237/2; 253/3; 278/1; 298/;
 C;Superfamily: glucose-6-phosphate isomerase
 C;Keywords: cytosol; gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomera

Query Match 94.1%; Score 32; DB 2; Length 560;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

```

|||||:
Db      227 TLREWI 232

RESULT 5
T09154
glucose-6-phosphate isomerase (EC 5.3.1.9), cytosol - spinach
C:Species: Spinacia oleracea (spinach)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09154
R:Nowitzki, U.; Flechner, A.; Kellermann, J.; Hasegawa, M.; Schnarrenberger, C.; Martin,
Gene 214, 205-213, 1998
A:Title: Eubacterial origin of nuclear genes for chloroplast and cytosolic glucose-6-ph
A:Reference number: 216593; MUID:98322268; PMID:9651529
A:Accession: T09154
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-566 <NOW>
A:CROSS-references: UNIPROT:O82059; UNIPARC:UPI000016DF8A; EMBL:AJ000266; NID:g3413512;
C:Genetics:
A:Gene: GPIS
C:Function:
A:Description: catalyzes reversible isomerization of glucose-6-phosphate and fructose-6-
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match      94.1%; Score 32; DB 2; Length 566;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TLREWL 6
      |||||:
Db      226 TLREWI 231

RESULT 6
S57830
glucose-6-phosphate isomerase (EC 5.3.1.9) isoenzyme 1, cytosolic - Clarkia xantiana
C:Species: Clarkia xantiana
C:Date: 28-Oct-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: S57830
R:Ford, V.S.; Thomas, B.R.; Gottlieb, L.D.
Syst. Bot. 20, 147-160, 1995
A:Title: The same duplication accounts for the PgiC genes in Clarkia xantiana and C. lew
A:Reference number: S57830
A:Accession: S57830
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-568 <FOR>
A:CROSS-references: UNIPROT:P54240; UNIPARC:UPI000012AEFE; EMBL:X80666; NID:g784973; PI
C:Genetics:
A:Gene: PgiC1
A:Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: cytosol; intramolecular oxidoreductase; isomerase

Query Match      94.1%; Score 32; DB 2; Length 568;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TLREWL 6
      |||||:
Db      226 TLREWI 231

RESULT 7
S23542
glucose-6-phosphate isomerase (EC 5.3.1.9) 2a, cytosolic - farewell-to-spring (Clarkia 1
N:Alternate names: phosphoglucose isomerase
C:Species: Clarkia lewisii
C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S23542

```

```

R;Thomas, B.R.; Laudencia-Chinguanco, D.; Gottlieb, L.D.
Plant Mol. Biol. 19, 745-757, 1992
A:Title: Molecular analysis of the plant gene encoding cytosolic phosphoglucose isomeras
A:Reference number: S23542; MUID:92353384; PMID:1643281
A:Accession: S23542
A:Molecule type: DNA
A:Residues: 1-569 <THO>
A:CROSS-references: UNIPROT:P29333; UNIPARC:UPI000012AF03; EMBL:X64332; NID:g18055; PIDN
C:Genetics:
A:Gene: pgiC2a
A:Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C:Complex: homodimer
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: cytosol; gluconeogenesis; glycolysis; homodimer; intramolecular oxidoreducta

Query Match      94.1%; Score 32; DB 2; Length 569;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TLREWL 6
      |||||:
Db      226 TLREWI 231

RESULT 8
S41806
glucose-6-phosphate isomerase (EC 5.3.1.9) 1a, cytosolic - farewell-to-spring (Clarkia 1
N:Alternate names: phosphoglucose isomerase
C:Species: Clarkia lewisii
C:Date: 20-May-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S41806
R;Thomas, B.R.; Ford, V.S.; Pichersky, E.; Gottlieb, L.D.
Genetics 135, 895-905, 1993
A:Title: Molecular characterization of duplicate cytosolic phosphoglucose isomerase gene
A:Reference number: S41806; MUID:94123980; PMID:8293986
A:Accession: S41806
A:Molecule type: DNA
A:Residues: 1-569 <THO>
A:CROSS-references: UNIPROT:P34796; UNIPARC:UPI000012AEFA; EMBL:X71084; NID:g415942; PI
C:Genetics:
A:Gene: pgiC1a
A:Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: cytosol; gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomera

Query Match      94.1%; Score 32; DB 2; Length 569;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TLREWL 6
      |||||:
Db      226 TLREWI 231

RESULT 9
S57831
glucose-6-phosphate isomerase (EC 5.3.1.9) isoenzyme 2, cytosolic - Clarkia xantiana
C:Species: Clarkia xantiana
C:Date: 28-Oct-1995 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C:Accession: S57831
R;Ford, V.S.; Thomas, B.R.; Gottlieb, L.D.
Syst. Bot. 20, 147-160, 1995
A:Title: The same duplication accounts for the PgiC genes in Clarkia xantiana and C. le
A:Reference number: S57830
A:Accession: S57831
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-569 <FOR>
A:CROSS-references: UNIPROT:P54242; UNIPARC:UPI000012AF04; EMBL:X80667; NID:g784975; PI
C:Genetics:
A:Gene: PgiC2
A:Introns: 17/3; 43/2; 70/3; 86/3; 138/3; 171/1; 185/2; 208/1; 236/2; 252/3; 277/1; 297/
C:Superfamily: glucose-6-phosphate isomerase

```

C;Keywords: cytosol; intramolecular oxidoreductase; isomerase

Query Match 94.1%; Score 32; DB 2; Length 569;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 226 TLREWI 231

RESULT 10

S60179
pol polyprotein homolog - fungus (Fusarium oxysporum) retrotransposon skippy
C;Species: Fusarium oxysporum
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
C;Accession: S60179
R;Anaya, N.; Roncero, M.I.G.
Mol. Gen. Genet. 249, 637-647, 1995
A;Title: Skippy, a retrotransposon from the fungal plant pathogen Fusarium oxysporum.
A;Reference number: S60178; MUID:96132549; PMID:8544829
A;Accession: S60179
A;Molecule type: DNA
A;Residues: 1-1295 <ANA>
A;Cross-references: UNIPARC:UPI000011DE25; EMBL:L34658; NID:G510695; PIDN:AAA88791.1; PID:gl6502079;
A;Experimental source: retrotransposon skippy
C;Genetics:

A;Mobile element: retrotransposon skippy
C;Superfamily: Fusarium retrotransposon skippy pol polyprotein; chromobox homology
C;Keywords: polyprotein
F/1236-1273/Domain: chromobox homology <CBH>

Query Match 94.1%; Score 32; DB 2; Length 1295;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 336 TLREWI 341

RESULT 11

F75291
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: F75291
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: F75291

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-85 <WHI>

A;Cross-references: UNIPROT:Q9RS35; UNIPARC:UPI000000CIAC3; GB:AE002061; GB:AE000513; NID

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR2292

A;Map position: 1

Query Match 91.2%; Score 31; DB 2; Length 85;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 26 TLQWL 31

RESULT 12

T38761
hypothetical protein SPAC3H8.03 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38761
R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, January 1996

A;Reference number: Z21810

A;Accession: T38761

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-105 <GEN>

A;Cross-references: UNIPROT:Q10139; UNIPARC:UPI000013A155; EMBL:Z69086; PIDN:CAA93160.1;

A;Experimental source: strain 972h-; cosmid c3H8

C;Genetics:

A;Gene: SPDB:SPAC3H8.03

A;Map position: 1

A;Introns: 60/3

Query Match 91.2%; Score 31; DB 2; Length 105;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 95 TVREWL 100

RESULT 13

AI0605
conserved hypothetical protein STY0909 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AI0605
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AI0605

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-158 <PAR>

A;Cross-references: UNIPARC:UPI000005A0B2; GB:AL513382; PIDN:CAD05315.1; PID:gl6502079;

C;Genetics:

A;Gene: STY0909

C;Superfamily: Escherichia coli hypothetical protein ybjN

Query Match 91.2%; Score 31; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||||
Db 11 TLQWL 16

RESULT 14

E64823
hypothetical protein ybjN - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: E64823
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E64823
A;Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-158 <BLAT>
A:Cross-references: UNIPROT:P75915; UNIPARC:UPI00000480CD; GB:AE000187; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: ybJn
C:Superfamily: Escherichia coli hypothetical protein ybJn

Query Match 91.2%; Score 31; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
DB 11 TLQWL 16

RESULT 15
E90745
Probable sensory transduction regulator [imported] - Escherichia coli (strain O157:H7, e
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E90745
R:Havashi, T.; Makino, K.; Kurokawa, K.; Ohnishi, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90745
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <HAY>
A:Cross-references: UNIPROT:P75915; UNIPARC:UPI00000480CD; GB:BA000007; PIDN:BA834356.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs0933
C:Superfamily: Escherichia coli hypothetical protein ybJn

Query Match 91.2%; Score 31; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
DB 11 TLQWL 16

RESULT 16
C83872
Hypothetical protein BH1779 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C83872
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C83872
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-287 <STO>
A:Cross-references: UNIPROT:Q9KB25; UNIPARC:UPI00000C3CC4; GB:AP001513; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1779
C:Superfamily: scyllo-inosamine-4-phosphate amidinotransferase

Query Match 91.2%; Score 31; DB 2; Length 287;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||

A:Molecule type: DNA
A:Residues: 1-299 <SIM>
A:Cross-references: UNIPARC:UPI0000165A87; GB:AE004068; GB:AE003849; NID:G9107832; PIDN:
A:Experimental source: strain 9a5C
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2609

Query Match 91.2%; Score 31; DB 2; Length 299;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||||
DB 158 TLQWL 163

RESULT 18
S34816
nitrogenase cofactor synthesis protein nifs - Rhodobacter capsulatus
N:Contains: L-cysteine sulfoxyltransferase (EC 2.8.1.1-)
C:Species: Rhodobacter capsulatus
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S34816
R:Masopohl, B.; Angermueller, S.; Hennecke, S.; Huebner, P.; Moreno-Vivian, C.; Klipp, W
Mol. Gen. Genet. 238, 369-382, 1993
A:Title: Nucleotide sequence and genetic analysis of the Rhodobacter capsulatus ORF6-nif
A:Reference number: S34814; MUID:93261420; PMID:8492805
A:Accession: S34816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <MAS>
A:Cross-references: UNIPROT:Q07177; UNIPARC:UPI0000130115; EMBL:X68444; NID:G297915; PII
C:Genetics:
A:Gene: nifs
C:Superfamily: nitrogen fixation protein nifs
C:Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase
P.203/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
F.325/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 384;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 114 TLKEWL 119

RESULT 17

B82535
conserved hypothetical protein XF2609 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82535
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82535
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <SIM>
A:Cross-references: UNIPARC:UPI0000165A87; GB:AE004068; GB:AE003849; NID:G9107832; PIDN:
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2609

Query Match 91.2%; Score 31; DB 2; Length 299;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6

|||||

DB 158 TLQWL 163

```

QY      1 TLREWL 6
      |||:||
Db      267 TLRDWL 272

RESULT 19
S01825
transforming growth factor beta-3 precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C;Accession: S01825
R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Ma
EMBO J. 7, 3737-3743, 1988
A;Title: A new type of transforming growth factor-beta, TGF-beta3.
A;Reference number: S01824; MUID:89091120; PMID:3208746
A;Accession: S01825
A;Molecule type: mRNA
A;Residues: 1-409 <DER>
A;Cross-references: UNIPROT:P15203; UNIPARC:UPI0000136CB9; EMBL:X14150; NID:g2127; PIDN:
C;Superfamily: inhibin
C;Keywords: growth factor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-297/Domain: propeptide #status predicted <PRO>
F;298-409/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match      91.2%; Score 31; DB 2; Length 409;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TLREWL 6
      |||:||
Db      205 TVREWL 210

RESULT 20
A41397
transforming growth factor beta-3 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
C;Accession: A41397; A61039; A61225
R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
Mol. Endocrinol. 3, 1926-1934, 1989
A;Title: Complementary DNA cloning of the murine transforming growth factor-beta3 (TGFB
tissues.
A;Reference number: A41397; MUID:90190650; PMID:2628730
A;Accession: A41397
A;Molecule type: mRNA
A;Residues: 1-410 <MTL>
A;Cross-references: UNIPROT:P17125; UNIPARC:UPI0000020D6C; GB:M32745; NID:g201949; PIDN:
R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.
Growth Factors 3, 139-146, 1990
A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-beta3.
A;Reference number: A61039; MUID:91000714; PMID:2206556
A;Accession: A61039
A;Molecule type: mRNA
A;Residues: 1-410 <DEN>
A;Cross-references: UNIPARC:UPI0000020D6C
R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.
Cell Growth Differ. 2, 77-83, 1991
A;Title: Cell lineage specificity of expression of the murine transforming growth factor
A;Reference number: A61225; MUID:91299576; PMID:2069871
A;Accession: A61225
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 285-410 <MAT>
A;Cross-references: UNIPARC:UPI0000176557
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; growth regulation
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-298/Domain: propeptide #status predicted <PRO>
F;259-261/Region: cell attachment (R-G-D) motif
F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>
F;772,133,140/Binding site: carboxydrate (Asn) (covalent) #status predicted
```

```

Query Match      91.2%; Score 31; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TLREWL 6
      |||:||
Db      205 TVREWL 210

RESULT 21
A55706
transforming growth factor beta-3 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 09-Jul-2004
C;Accession: A55706; B40699; S36042
R;Wang, J.; Kuliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Teeu, I.; Post, M.
J. Biol. Chem. 270, 2722-2728, 1995
A;Title: Cloning and expression of glucocorticoid-induced genes in fetal rat lung fibrobl
A;Reference number: A55706; MUID:95155340; PMID:7852342
A;Accession: A55706
A;Molecule type: mRNA
A;Residues: 1-410 <WAN>
A;Cross-references: UNIPROT:Q07258; UNIPARC:UPI000017655A; GB:U03491
A;Note: it is uncertain whether Met-1 is the initiator
R;McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.
J. Cell Biol. 121, 1397-1407, 1993
A;Title: A role for TGF-beta in oligodendrocyte differentiation.
A;Reference number: A40699; MUID:93286190; PMID:8509457
A;Accession: B40699
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 157-211 <MCK>
A;Cross-references: UNIPARC:UPI0000170B3A; EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PI
C;Superfamily: inhibin

Query Match      91.2%; Score 31; DB 2; Length 410;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 TLREWL 6
      |||:||
Db      205 TVREWL 210

RESULT 22
A34939
transforming growth factor beta-3 precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C;Accession: A34939; S25850; S36125; S36124; I51181
R;Jakowlew, S.B.; Dillard, P.J.; Kondaiah, P.; Sporn, M.B.; Roberts, A.B.
Mol. Endocrinol. 2, 747-755, 1988
A;Title: Complementary deoxyribonucleic acid cloning of a novel transforming growth factor
A;Reference number: A34939; MUID:89096966; PMID:3211158
A;Accession: A34939
A;Status: preliminary; not compared with conceptual translation
A;Residues: 1-412 <JAK>
A;Cross-references: UNIPROT:P16047; UNIPARC:UPI00001713BB; GB:M31154; NID:g212758; PIDN:
R;Burt, D.W.; Paton, I.R.; Dey, B.R.
J. Mol. Endocrinol. 7, 175-183, 1991
A;Title: Comparative analysis of human and chicken transforming growth factor-beta-2 and
A;Reference number: S25850; MUID:92134496; PMID:1840616
A;Accession: S25850
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-117 <BUR>
A;Cross-references: UNIPARC:UPI000017146B; EMBL:X58127; NID:g63815; PIDN:CAA41128.1; PID
A;Accession: S36125
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 119-172 <BU2>
```


C;Cross-references: UNIPARC:UPI0000176560; EMBL:X60055; NID:g396688; PIDN:CAA42653.1; PID:1-27/Domain: signal sequence #status predicted <SIG>
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Accession: S36124
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 173-322, 'ELPT', 327-412 <BU3>
A;Cross-references: UNIPARC:UPI0000176561; EMBL:X60091
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
R;Jakowlew, S.B.; Lechleider, R.; Geiser, A.G.; Kim, S.J.; Santa-Coloma, T.A.; Cubert, J.
Mol. Endocrinol. 6, 1285-1298, 1992
A;Title: Identification and characterization of the chicken transforming growth factor-beta
A;Reference number: I51181; MUID:93024487; PMID:1406706
A;Accession: I51181
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-117 <J2>
A;Cross-references: UNIPARC:UPI000017146B; GB:S46000; NID:g257172; PIDN:AAB23575.1; PID:1-22/Domain: signal sequence #status predicted <SIG>
C;Genetics:
A;Introns: 216/1; 252/1; 309/2; 360/3
A;Note: list of introns may be incomplete
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-300/Domain: propeptide #status predicted <PRO>
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>
F;74,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 412;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
DB 207 TVREWL 212

RESULT 23
A36169
Transforming growth factor beta-3 precursor - human
C;Species: Homo sapiens (man)
C;Date: 14-Dec-1990 #sequence revision 14-Dec-1990 #text_change 09-Jul-2004
C;Accession: A36169; A41262; S01824
R;ten Dijke, P.; Hansen, P.; Iwata, K.K.; Pieler, C.; Foulkes, J.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 4715-4719, 1988
A;Title: Identification of another member of the transforming growth factor type beta gene
A;Reference number: A36169; MUID:88263019; PMID:3164476
A;Accession: A36169
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-412 <TENS>
A;Cross-references: UNIPROT:P10600; UNIPARC:UPI000000D8BC; GB:J03241; NID:g339551; PIDN:R;Arrick, B.A.; Lee, A.L.; Grendell, R.L.; Derynck, R.
Mol. Cell. Biol. 11, 4306-4313, 1991
A;Title: Inhibition of translation of transforming growth factor-beta3 mRNA by its 5' untranslated region
A;Reference number: A41262; MUID:91342629; PMID:1875922
A;Accession: A41262
A;Molecule type: DNA
A;Residues: 1-48 <ARR>
A;Cross-references: UNIPARC:UPI000017655E; GB:M58524
R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee, L.; Mas
EMBO J. 7, 3737-3743, 1988
A;Title: A new type of transforming growth factor-beta, TGF-beta3.
A;Reference number: S01824; MUID:89091120; PMID:3208746
A;Accession: S01824
A;Molecule type: mRNA
A;Residues: 3-412 <DER>
A;Cross-references: UNIPARC:UPI000016B0BF; EMBL:X14885; NID:g37075; PIDN:CAA33024.1; PID:1-27/Domain: signal sequence #status predicted <SIG>
C;Genetics:
A;Gene: GDB:TGF3
A;Cross-references: GDB:120437; OMIM:190230
A;Map position: 14q24-14q24
A;Introns: 118/1; 172/3; 216/1; 252/1; 309/2; 360/3
C;Superfamily: inhibin

C;Keywords: growth factor; homodimer
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-300/Domain: propeptide #status predicted <PRO>
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 91.2%; Score 31; DB 2; Length 412;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
DB 207 TVREWL 212

RESULT 24
O4PSCP
camphor 5-monooxygenase (EC 1.14.15.1) cytochrome P450 101 - Pseudomonas putida plasmid
N;Alternate names: cytochrome P450-CAM
C;Species: Pseudomonas putida
C;Date: 30-Apr-1982 #sequence_revision 31-Dec-1993 #text_change 05-Oct-2004
C;Accession: A25660; S34614; C60886; A00194
R;Unger, B.P.; Gunsalus, I.C.; Sligar, S.G.
J. Biol. Chem. 261, 1158-1163, 1986
A;Title: Nucleotide sequence of the Pseudomonas putida cytochrome P-450-cam gene and its
A;Reference number: A94678; MUID:86111751; PMID:3003058
A;Accession: A25660
A;Molecule type: DNA
A;Residues: 1-415 <UNG>
A;Cross-references: UNIPROT:P00183; UNIPARC:UPI00001118C2; GB:M12546; NID:g151114; PIDN:R;Aramaki, H.; Koga, H.; Sagara, Y.; Hosoi, M.; Horiuchi, T.
Biochim. Biophys. Acta 1174, 91-94, 1993
A;Title: Complete nucleotide sequence of the 5-exo-hydroxycamphor dehydrogenase gene on
A;Reference number: S34613; MUID:93326643; PMID:8334169
A;Accession: S34614
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-42 <ARA>
A;Cross-references: UNIPARC:UPI0000171DDF
A;Experimental source: PpG1; ATCC 17453; CAM plasmid
R;Romeo, C.; Moriaki, N.; Yasunobu, K.T.; Gunsalus, I.C.; Koga, H.
J. Protein Chem. 6, 253-261, 1987
A;Title: Identification of the coding region for the putidaredoxin reductase gene from
A;Reference number: A60886
A;Accession: C60886
A;Molecule type: DNA
A;Residues: 408-415 <ROM>
A;Cross-references: UNIPARC:UPI0000171DE0
R;Hanlu, M.; Ames, L.G.; Yasunobu, K.T.; Shastri, B.A.; Gunsalus, I.C.
J. Biol. Chem. 257, 12664-12671, 1982
A;Title: Amino acid sequence of the Pseudomonas putida cytochrome P-450. II. Cyanogen
A;Reference number: A00194; MUID:83030788; PMID:7130171
A;Accession: A00194
A;Molecule type: protein
A;Residues: 2-55,58-276, 'Q', 278-361, 'S', 363-407, 'N', 409-415 <HAN>
A;Cross-references: UNIPARC:UPI0000171DE1
C;Genetics:
A;Gene: camC; CYP101
A;Genome: plasmid
C;Function:
A;Description: catalyzes hydroxylation of camphor to yield 5-exo-hydroxycamphor; electri
A;Superfamily: camphor 5-monooxygenase; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase,
F;246-380/Domain: cytochrome P450 homology <CYP>
F;358/Binding site: heme iron (Cys) (axial ligand) #status experimental

Query Match 91.2%; Score 31; DB 1; Length 415;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|:||||
DB 371 TLKWL 376

RESULT 25
T19971
hypothetical protein C46F11.5a - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19971
R/Burton, J.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19206
A/Accession: T19971
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-505 <WTL>
A/Cross-references: UNIPROT:Q93383; UNIPARC:UPI0000179966; EMBL:Z81449; PIDN:CAB03762.1;
A/Experimental source: clone C46F11
C/Genetics:
A/Gene: CESP:C46F11.5a
A/Map position: 3
A/Introns: 38/2; 89/3; 119/1; 178/1; 195/1; 260/3; 313/2; 366/3; 407/1; 450/1; 480/3

Query Match 91.2%; Score 31; DB 2; Length 505;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 252 TLKEWL 257

RESULT 26
T19973
hypothetical protein C46F11.5b - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T19973
R/Burton, J.
submitted to the EMBL Data Library, November 1996
A/Reference number: Z19206
A/Accession: T19973
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-506 <WTL>
A/Cross-references: UNIPROT:O62100; UNIPARC:UPI0000179965; EMBL:Z81449; PIDN:CAB03764.1;
A/Experimental source: clone C46F11
C/Genetics:
A/Gene: CESP:C46F11.5b
A/Map position: 3
A/Introns: 38/2; 89/3; 119/1; 178/1; 195/1; 260/3; 313/2; 366/3; 407/1; 450/1; 481/3

Query Match 91.2%; Score 31; DB 2; Length 506;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 252 TLKEWL 257

RESULT 27
A96018
Probable regulatory protein, possibly two-component response regulator SMB20702 [importe
C/Species: *Sinorhizobium meliloti*
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: A96018
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1.683-kb pSymA megaplasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: A96018
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-554 <KUR>

A/Cross-references: UNIPROT:Q92TT6; UNIPARC:UPI00000CB867; GB:AL591985; PIDN:CAC49809.1;
A/Experimental source: strain 1021, megaplasmid pSymB
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;
A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: SMB20702
A/Genome: plasmid

Query Match 91.2%; Score 31; DB 2; Length 554;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
|||
Db 527 TLRDWL 532

RESULT 28
A56126
peroxisomal targeting signal 1 receptor - human
N/Alternate names: peroxisomal C-terminal targeting signal 1 import receptor; peroxisomal
C/Species: *Homo sapiens* (man)
C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 31-Dec-2004
C/Accession: A56126; I38142; I38925; S53100
R/Fransen, M.; Brees, C.; Baumgart, E.; Vanhooren, J.C.T.; Baes, M.; Mannaerts, G.P.; Var
J. Biol. Chem. 270, 7731-7736, 1995
A/Title: Identification and characterization of the putative human peroxisomal C-terminal
A/Reference number: A56126; MUID:95221441; PMID:7706321
A/Accession: A56126
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-639 <FRA>
A/Cross-references: UNIPROT:P50542; UNIPARC:UPI000016AE82; GB:X84899; NID:g695565; PIDN:C
R/Wiemer, E.A.; Nuttley, W.M.; Bertolaet, B.L.; Li, X.; Francke, U.; Wheelock, M.J.; Am
J. Cell Biol. 130, 51-65, 1995
Nature Genet. 9, 115-125, 1995
A/Title: Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the
A/Reference number: A56735; MUID:95310365; PMID:7790377
A/Accession: I38142
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-214,252-639 <WIE>
A/Cross-references: UNIPARC:UPI00001316E7; EMBL:Z48054; NID:9732797; PIDN:CAA88131.1; PFI
R/Dott, G.; Braverman, N.; Wong, C.; Moser, A.; Moser, H.W.; Watkins, P.; Valle, D.; Gou
A/Title: Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the
A/Reference number: I38925; MUID:95235555; PMID:7719337
A/Accession: I38925
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-214,252-424,'I', 426-639 <DOD>
A/Cross-references: UNIPARC:UPI000016A16C; EMBL:U19721; NID:g694004; PIDN:AAC50103.1; PFI
C/Genetics:
A/Gene: GDB:PXRL1; PEX5; PTS1R
A/Cross-references: GDB:433739; OMIM:600414
A/Map position: 12p13-12p13
C/Function:
A/Description: binds the carboxyl-terminal tripeptide Ser-[Lys/Arg/His]-Leu, SKL, or per
C/Keywords: peroxisome biogenesis; protein transport
F;452-485/Domain: tetratricopeptide repeat homology <TT1>
F;488-521/Domain: tetratricopeptide repeat homology <TT2>
F;522-555/Domain: tetratricopeptide repeat homology <TT3>

Query Match 91.2%; Score 31; DB 2; Length 639;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||:|
Db 425 TLRDNL 430

RESULT 29
E96508
hypothetical protein T12C22.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E96508
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96508
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-936 <STO>
A:Cross-references: UNIPROT:Q9LDP9; UNIPARC:UPI00000AA9E4; GB:AE005173; NID:G8656002; PI
C:Genetics:
A:Gene: T12C22.19
A:Map position: 1
C:Superfamily: replication licensing factor MCM2; MCM homology

Query Match 91.2%; Score 31; DB 2; Length 936;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||:|
Db 204 TLREWL 209

RESULT 30
T48216
hypothetical protein T20L15.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 31-Dec-2004
C:Accession: T48216
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkee, W.; Stiekema, W.; Bancroft, I.; New
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24488
A:Accession: T48216
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-984 <BEV>
A:Cross-references: UNIPROT:Q9LZV1; UNIPARC:UPI0000048AAP; EMBL:AL162351
A:Experimental source: cultivar Columbia; BAC clone T20L15
C:Genetics:
A:Map position: 5
A:Introns: 29/1; 164/2; 212/2; 236/2; 260/2; 285/2; 308/2; 331/2; 355/2; 384/2; 407/2; 5
A:Note: T20L15.220
C:Superfamily: Receptor-like protein kinase

Query Match 91.2%; Score 31; DB 2; Length 984;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||:|
Db 739 TLRDNL 744

RESULT 31
F70697

probable arabinosyltransferase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70697
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70697
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1094 <COL>
A:Cross-references: UNIPROT:P72060; UNIPARC:UPI0000129594; GB:Z80343; GB:AL123456; NID:G
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: emba
C:Superfamily: probable arabinosyl transferase

Query Match 91.2%; Score 31; DB 2; Length 1094;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||:|
Db 232 TLRDNL 237

RESULT 32
T00390
KIAA0614 protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00390; T17254
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00390
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1630 <ISH>
A:Cross-references: UNIPROT:Q9UFT6; UNIPARC:UPI000003B446; EMBL:AB014514; NID:G3327041;
R:Kocher, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17254
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1515-1630 <KOB>
A:Cross-references: UNIPARC:UPI0000073F07; EMBL:AL117469
A:Experimental source: adult uterus; clone DKFZp58601022
C:Genetics:
A:Note: KIAA0614; DKFZp58601022.1

Query Match 91.2%; Score 31; DB 2; Length 1630;
Best Local Similarity 83.3%; Pred. No. 8.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|||:|
Db 165 TVREWL 170

RESULT 33
H83343
probable non-ribosomal peptide synthetase PA2424 [imported] - Pseudomonas aeruginosa (s
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83343
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: H83343
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-4342 <STO>
A/Cross-references: UNIPROT:Q91157; UNIPARC:UPI0000110218; GB:AE004669; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA2424
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F/60-553/Domain: acetate-CoA ligase homology <ACLI>
F/584-652/Domain: acyl carrier protein homology <ACP1>
F/1174-1622/Domain: acetate-CoA ligase homology <ACL2>
F/1637-1705/Domain: acyl carrier protein homology <ACP2>
F/2132-2689/Domain: acetate-CoA ligase homology <ACL3>
F/2706-2773/Domain: acyl carrier protein homology <ACP3>
F/3759-4230/Domain: acetate-CoA ligase homology <ACL4>
F/4248-4316/Domain: acyl carrier protein homology <ACP4>
F/1669,2738,4280/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 2; Length 4342;
Best Local Similarity 83.3%; Pred. No. 2.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|:||||
Db 3560 TVREWL 3565

RESULT 34
S10865
early E4 13K protein - human adenovirus 12
C/Species: Mastadenovirus h12 (human adenovirus 12)
C/Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: S10865; S33953
R/Hogenkamp, T.; Esche, H.
Nucleic Acids Res. 18, 3065-3066, 1990
A/Title: Nucleotide sequence of the right 10% of adenovirus type 12 DNA encoding the ent
A/Reference number: S10860; MUID:90272430; PMID:2349112
A/Accession: S10865
A/Molecule type: DNA
A/Residues: 1-120 <HOG>
A/Cross-references: UNIPROT:P36709; UNIPARC:UPI000003654C; EMBL:X51800
A/Note: the authors translated the codon TGG for residue 83 as Tyr
R/Sprengel, J.
submitted to the EMBL Data Library, June 1993
A/Reference number: S33928
A/Accession: S33953
A/Molecule type: DNA
A/Residues: 1-120 <SPR>
A/Cross-references: UNIPARC:UPI000003654C; EMBL:X73487; NID:G313361; PIDN:CAA51902.1; PI
C/Superfamily: adenovirus early E4 13K protein

Query Match 88.2%; Score 30; DB 2; Length 120;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|:||||
Db 53 SLREWL 58

RESULT 35
D72305
hypothetical protein - *Thermotoga maritima* (strain MSB8)
C/Species: *Thermotoga maritima*
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: D72305
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A/Reference number: A72200; MUID:99287316; PMID:10360571
A/Accession: D72305
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-158 <ARN>
A/Cross-references: UNIPROT:Q9X0A5; UNIPARC:UPI000000C1322; GB:AE001763; GB:AE000512; NID:
A/Experimental source: strain MSB8
C/Genetics:
A/Gene: TM1012

Query Match 88.2%; Score 30; DB 2; Length 158;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|:||||
Db 148 TLKWL 153

RESULT 36
G89799
hypothetical protein SA0330 [imported] - *Staphylococcus aureus* (strain N315)
C/Species: *Staphylococcus aureus*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 31-Dec-2004
C/Accession: G89799
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; i
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsesu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: G89799
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-181 <KUR>
A/Cross-references: UNIPROT:Q99WN5; UNIPARC:UPI000000CACCA; GB:BA000018; PID:gl3700256; P
A/Experimental source: strain N315
C/Genetics:
A/Gene: SA0330
C/Superfamily: ribosomal protein serine N-acetyltransferase

Query Match 88.2%; Score 30; DB 2; Length 181;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
|:||||
Db 33 SLREWL 38

RESULT 37
S52461
hypothetical protein 233 - *Coxiella burnetii*
C/Species: *Coxiella burnetii*
C/Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: S52461
R/Willems, H.; Thiele, D.; Valkova, D.
submitted to the EMBL Data Library, February 1995
A/Reference number: S52460
A/Accession: S52461
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-233 <WIL>
A/Cross-references: UNIPROT:Q45912; UNIPARC:UPI000000B946E; EMBL:X84722; NID:G682754; PID:

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 38
S38228
hypothetical protein - Coxiella burnetii
C:Species: Coxiella burnetii
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S38228
R:Thiele, D.; Willems, H.; Haas, M.; Krauss, H.
submitted to the EMBL Data Library, October 1993
A:Reference number: S38215
A:Accession: S38228
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <TH>
A:Cross-references: UNIPROT:Q45935; UNIPARC:UPI00000BBB8C; EMBL:X75356; NID:g407370; PID

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 39
S53104
hypothetical protein 233 - Coxiella burnetii
C:Species: Coxiella burnetii
C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: S53104
R:Ritter, M.; Thiele, D.; Willems, H.
submitted to the EMBL Data Library, March 1995
A:Reference number: S53104
A:Accession: S53104
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <RIT>
A:Cross-references: UNIPROT:Q45912; UNIPROT:Q45935; UNIPROT:O52877; UNIPARC:UPI000017970

Query Match 88.2%; Score 30; DB 2; Length 233;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 23 TLQEWL 28

RESULT 40
E69189
hypothetical protein MTH672 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: E69189
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwnani, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcu
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: E69189
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-239 <MTH>
A:Cross-references: UNIPROT:O26768; UNIPARC:UPI0000062B47; GB:AE000846; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH672

A:Start codon: GTG

Query Match 88.2%; Score 30; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 5
|||||
Db 91 TLREW 95

RESULT 41
A30227
hypothetical protein 2BE2121 - Chinese hamster (fragment)
C:Species: Cricetus griseus (Chinese hamster)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 01-Dec-2000
C:Accession: A30227; B30227
R:Foreman, P.K.; Hamlin, J.L.
Mol. Cell. Biol. 9, 1137-1147, 1989
A:Title: Identification and characterization of a gene that is coamplified with dihydrof
A:Reference number: A30227; MUID:89261717; PMID:2725490
A:Accession: A30227
A:Molecule type: mRNA
A:Residues: 1-263 <FOR>
A:Cross-references: UNIPARC:UPI00000BE6589; GB:M23159; NID:g191006; PIDN:AAA36962.1; PID:
A:Note: the source is Chinese hamster ovary cells
A:Accession: B30227
A:Molecule type: mRNA
A:Residues: 215-263 <PO2>
A:Cross-references: UNIPARC:UPI0000179949; GB:M23160; NID:g191008
A:Note: this ORF is not annotated in GenBank entry CIUDHFRCA, release 109.0
A:Note: the source is Chinese hamster ovary cells

Query Match 88.2%; Score 30; DB 2; Length 263;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
||:||||
Db 76 TLQEWL 81

RESULT 42
F83042
hypothetical protein PA4834 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83042
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83042
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-284 <STO>
A:Cross-references: UNIPROT:Q9HUX6; UNIPARC:UPI00000CSE2F; GB:AE004896; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4834

Query Match 88.2%; Score 30; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREW 5
|||||
Db 204 TLREW 208

RESULT 43

F86431
 hypothetical protein T518.8 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 A/Accession: F86431
 R/Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 ansen, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: F86431
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-292 <STO>
 A/Cross-references: UNIPROT:Q9SA78; UNIPARC:UPI000009F92E; GB:AE005172; NID:G4587519; PI
 C/Genetics:
 A/Map position: 1

Query Match 88.2%; Score 30; DB 2; Length 292;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 :|||||
 Db 93 SLREWL 98

RESULT 44

S77938
 EBNA-LP protein - human herpesvirus 4
 C/Species: human herpesvirus 4, Epstein-Barr virus
 C/Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 05-Oct-2004
 A/Accession: S77938; S42441
 R/Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
 submitted to the EMBL Data Library, September 1987
 A/Reference number: S77938
 A/Accession: S77938
 A/Molecule type: mRNA
 A/Residues: 1-308 <SAM>
 A/Cross-references: UNIPROT:Q69136; UNIPARC:UPI00000F09AD; EMBL:M13940; NID:G330401; PID
 R/Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5096-5100, 1986
 A/Title: Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a p
 A/Reference number: S42440; MUID:86259739; PMID:3460083
 A/Accession: S42441
 A/Molecule type: mRNA
 A/Residues: 'M', 134-308 <SAW>
 A/Cross-references: UNIPARC:UPI0000177CC9; EMBL:M13940
 C/Keywords: alternative splicing

Query Match 88.2%; Score 30; DB 2; Length 308;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 :|||||
 Db 275 SLREWL 280

RESULT 45

F82601
 5'-nucleotidase XF2089 [imported] - Xylella fastidiosa (strain 9a5c)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 A/Accession: F82601
 R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: F82601
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-319 <STM>
 A/Cross-references: UNIPROT:Q9PBQ1; UNIPARC:UPI00000C2928; GB:AE004024; GB:AE003849; NID
 A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H.
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000

A/Authors: Ferreira, V.C.A.; Perro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF2089

Query Match 88.2%; Score 30; DB 2; Length 319;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 5
 :|||||
 Db 257 TLREW 261

RESULT 46

A29711
 deacetoxycephalosporin C synthetase - fungus (Acremonium sp.)
 N/Alternate names: expandase; hydroxylase
 C/Species: Acremonium sp.
 C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 31-Dec-2004
 A/Accession: A29711; A41864
 R/Samson, S.M.; Dotzlauf, J.E.; Slisz, M.L.; Becker, G.W.; Van Frank, R.M.; Veal, L.E.; Ye
 Bio/Technology 5, 1207-1214, 1987
 A/Title: Cloning and expression of the fungal expandase/hydroxylase gene involved in cep
 A/Reference number: A29711
 A/Accession: A29711
 A/Status: not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-332 <SAM>
 A/Cross-references: UNIPROT:P11935; UNIPARC:UPI0000044675
 A/Note: the source is designated as Cephalosporium acremonium
 R/Gutierrez, S.; Velasco, J.; Fernandez, F.J.; Martin, J.P.
 J. Bacteriol. 174, 3056-3064, 1992
 A/Title: The cefG gene of Cephalosporium acremonium is linked to the cefEF gene and enco
 A/Reference number: A41864; MUID:92234966; PMID:1569032
 A/Accession: A41864
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-76, 'L', <GUT>
 A/Cross-references: UNIPARC:UPI0000175256
 A/Experimental source: strain C10
 A/Note: sequence extracted from NCBI backbone (NCBIN:104773, NCBI:P:97574); this ORF is n
 C/Superfamily: isopenicillin N synthase

Query Match 88.2%; Score 30; DB 2; Length 332;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 :|||||
 Db 294 TLREWL 299

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2963

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-394 <KUR>

A;Cross-references: UNIPROT:Q8UAR2; UNIPARC:UPI00000D2161; GB:AE008689; PIDN:RAL44124.1;

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3311

A;Map position: linear chromosome

Query Match 88.2%; Score 30; DB 2; Length 394;

Best Local Similarity 83.3%; Pred. No. 3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6

|||||

Db 18 TLRKWL 23

RESULT 50

G98319

mannose-6-phosphate isomerase (phosphomannose isomerase) (pmi) (phosphohexomutase) [impo

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C;Accession: G98319

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, B.;

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: G98319

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-394 <KUR>

A;Cross-references: UNIPROT:Q8UAR2; UNIPARC:UPI00000D2161; GB:AE007870; PIDN:AAK90081.1;

C;Genetics:

A;Gene: AGR_L3017

A;Map position: linear chromosome

Query Match 88.2%; Score 30; DB 2; Length 394;

Best Local Similarity 83.3%; Pred. No. 3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6

|||||

Db 18 TLRKWL 23

Search completed: May 12, 2006, 10:52:12

Job time : 15.3821 secs

RESULT 47

E90998

hypothetical protein ECe2957 [imported] - Escherichia coli (strain O157:H7, substrain R1

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: E90998

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.;

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: E90998

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-338 <HAY>

A;Cross-references: UNIPROT:Q9EYD4; UNIPARC:UPI00000B38DF; GB:BA0000007; PIDN:BA836380.1;

A;Experimental source: strain O157:H7, substrain R1MD 0509952

C;Genetics:

A;Gene: ECe2957

Query Match 88.2%; Score 30; DB 2; Length 338;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 5

|||||

Db 20 TLREW 24

RESULT 48

D85818

unknown protein encoded within prophage CP-933U [imported] - Escherichia coli (strain O1

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: D85818

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: D85818

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-346 <STO>

A;Cross-references: UNIPROT:Q8X8Z2; UNIPARC:UPI000016582A; GB:AE005174; NID:gi2516109; E

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: Z3092

Query Match 88.2%; Score 30; DB 2; Length 346;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREW 5

|||||

Db 28 TLREW 32

RESULT 49

AF2963

phosphomannose isomerase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AF2963

R;Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

i Karp, P.; Romero, P.; Zhang, S.;

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 42.4615 seconds
(without alignments)

99.694 Million cell updates/sec

Title: US-10-632-388-34

Perfect score: 34

Sequence: 1 TLREWL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trenbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	78	Q8Z85 PYRAE	Q8Z85 pyroaculum
2	34	100.0	78	Q50W5 ENTMOE	Q50W5 entamoeba h
3	34	100.0	190	Q9SH3 ARATH	Q9SH3 arabidopsis
4	34	100.0	213	Q4T3D3 TETNG	Q4T3D3 tetraodon n
5	34	100.0	243	Q7TNX8 RAT	Q7TNX8 rattus norv
6	34	100.0	258	LRP16 RAT	LRP16 rat
7	34	100.0	265	KSGA_KZOSF	KSGA_kzose
8	34	100.0	277	Q87IG5_VIBPA	Q87IG5 vibrio para
9	34	100.0	296	Q9KK21 OLICA	Q9KK21 oligotropha
10	34	100.0	325	LRP16 HUMAN	LRP16 homo sapien
11	34	100.0	391	Q7SPR0_PHYPA	Q7SPR0 physcomitre
12	34	100.0	391	Q7PSJ0_ANOGA	Q7PSJ0 anopheles g
13	34	100.0	473	Q9SGQ0_ARATH	Q9SGQ0 arabidopsis
14	34	100.0	522	Q5SGN6_EMENT	Q5SGN6 aspergillus
15	34	100.0	556	Q6XPC4_PHYIN	Q6XPC4 phytophthor
16	34	100.0	557	Q4HCY6_9DEIO	Q4HCY6 deinococcus
17	34	100.0	597	Q5W7M1_XENTR	Q5W7M1 xenopus tro
18	34	100.0	597	Q6FLE9_CANGA	Q6FLE9 candida gla
19	34	100.0	637	Q9VYJ0_DROME	Q9VYJ0 drosophila
20	34	100.0	963	Q9UVC1_CLAFU	Q9UVC1 cladoporiu
21	34	100.0	1243	Q923T8_PSESM	Q923T8 pseudomonas
22	34	100.0	2066	Q87W69_PSESM	Q87W69 pseudomonas
23	34	100.0	2066	Q6JD66_TOXGO	Q6JD66 toxoplasma
24	34	100.0	5072	Q6FELS_ACIAD	Q6FELS acinetobact
25	34	94.1	81	Q70SL4_9MYRT	Q70SL4 hauya heyde
26	34	94.1	134	Q70SL4_9MYRT	Q70SL4 circaea cor
27	34	94.1	134	Q70SL7_9MYRT	Q70SL7 fuchsia cyl
28	34	94.1	144	Q656A5_ORYSA	Q656A5 oryza sativ
29	34	94.1	176	Q8D9L1_VIBVU	Q8D9L1 vibrio vuln
30	34	94.1	195	Q6N1X5_RHOPA	Q6N1X5 rhodopseudo
31	34	94.1	209	Q6N1X5_RHOPA	Q6N1X5 rhodopseudo
32	34	94.1	209	Q70SK1_LOPEZ	Q70SK1 lopezia gra
33	34	94.1	269	Q70SK1_9MYRT	Q70SK1 clarkia mil

32	94.1	281	2	Q70SK3_9MYRT	Q70SK3 megacorax g
33	94.1	307	2	Q8RW74_CLAEP	Q8RW74 clarkia epi
34	94.1	313	1	TOC34 ARATH	TOC34 arabidopsis
35	94.1	328	2	Q5V5F2_HALMA	Q5V5F2 haloarcula
36	94.1	342	2	Q6MB42_PARUM	Q6MB42 parachlamyd
37	94.1	357	2	O22092_9LILI	O22092 dioscorea n
38	94.1	357	2	O22093_9LILI	O22093 dioscorea q
39	94.1	357	2	O22094_9LILI	O22094 dioscorea s
40	94.1	357	2	O22095_9LILI	O22095 dioscorea t
41	94.1	357	2	O24644_9LILI	O24644 dioscorea g
42	94.1	368	2	O4SIN4_TETNG	O4SIN4 tetraodon n
43	94.1	371	2	Q70SK2_9MYRT	Q70SK2 hauya elega
44	94.1	371	2	Q70SK7_9MYRT	Q70SK7 epilobium c
45	94.1	373	2	Q70SK5_9MYRT	Q70SK5 epilobium b
46	94.1	380	2	Q8L6V2_9MYRT	Q8L6V2 clarkia dud
47	94.1	380	2	Q8L6V3_9MYRT	Q8L6V3 clarkia dud
48	94.1	380	2	Q8L6V4_9MYRT	Q8L6V4 clarkia dud
49	94.1	380	2	Q8RW61_CLAUN	Q8RW61 clarkia ung
50	94.1	381	2	Q70SL1_LUDPE	Q70SL1 ludwigia pe
51	94.1	402	2	Q6DIP8_XENTR	Q6DIP8 xenopus tro
52	94.1	411	2	O23903_9LILI	O23903 dioscorea t
53	94.1	429	2	Q8RW83_9MYRT	Q8RW83 clarkia del
54	94.1	436	2	O22078_9LILI	O22078 dioscorea q
55	94.1	443	2	Q9P858_PHANO	Q9P858 phaeosphaer
56	94.1	450	2	Q9SLB9_ARATH	Q9SLB9 arabidopsis
57	94.1	476	2	Q8VXR4_9MYRT	Q8VXR4 clarkia del
58	94.1	491	2	Q70SK9_GAULI	Q70SK9 gaura lindh
59	94.1	499	2	Q8L6U8_9MYRT	Q8L6U8 clarkia lin
60	94.1	499	2	Q8L6U9_9MYRT	Q8L6U9 clarkia lin
61	94.1	501	2	Q70SL0_9MYRT	Q70SL0 calylophus
62	94.1	501	2	Q8RW62_CLAUN	Q8RW62 clarkia ung
63	94.1	501	2	Q8RW72_9MYRT	Q8RW72 clarkia mod
64	94.1	516	2	Q8RW64_9MYRT	Q8RW64 clarkia sim
65	94.1	526	2	Q8RW65_9MYRT	Q8RW65 clarkia sim
66	94.1	526	2	Q8RW66_9MYRT	Q8RW66 clarkia sim
67	94.1	526	2	Q8RW79_9MYRT	Q8RW79 clarkia del
68	94.1	527	2	Q8RU88_9MYRT	Q8RU88 clarkia sim
69	94.1	527	2	Q8RW63_9MYRT	Q8RW63 clarkia sim
70	94.1	527	2	Q8RW76_9MYRT	Q8RW76 clarkia del
71	94.1	527	2	Q8RW81_9MYRT	Q8RW81 clarkia del
72	94.1	529	2	Q839B8_ENTFA	Q839B8 enterococcu
73	94.1	534	2	Q88Z37_LACPL	Q88Z37 lactobacill
74	94.1	559	2	Q8L6V0_9MYRT	Q8L6V0 clarkia het
75	94.1	560	1	G6PI_ARAGE	G6PI arabis gemm
76	94.1	560	1	G6PI_ARALP	G6PI arabidopsis
77	94.1	560	1	G6PI_ARATH	G6PI arabidopsis
78	94.1	560	2	Q7G121_9BRAS	Q7G121 leavenworth
79	94.1	560	2	Q7X833_ARAGE	Q7X833 arabis gemm
80	94.1	560	2	Q7X8A0_ARAGE	Q7X8A0 arabis gemm
81	94.1	560	2	Q7X9W9_ARAGE	Q7X9W9 arabis gemm
82	94.1	560	2	Q7X9X0_ARAGE	Q7X9X0 arabis gemm
83	94.1	560	2	Q7X9X1_ARAGE	Q7X9X1 arabis gemm
84	94.1	560	2	Q7X9X2_ARAGE	Q7X9X2 arabis gemm
85	94.1	560	2	Q7X9X3_ARAGE	Q7X9X3 arabis gemm
86	94.1	560	2	Q7X9X4_ARAGE	Q7X9X4 arabis gemm
87	94.1	560	2	Q7X9X5_ARAGE	Q7X9X5 arabis gemm
88	94.1	560	2	Q7X9X6_ARAGE	Q7X9X6 arabis gemm
89	94.1	560	2	Q8H9A4_ARAGE	Q8H9A4 arabis gemm
90	94.1	560	2	Q8H9A5_9BRAS	Q8H9A5 crucihimala
91	94.1	560	2	Q8H9A6_ARAGL	Q8H9A6 arabis glab
92	94.1	560	2	Q94JTI_ARATH	Q94JTI arabidopsis
93	94.1	560	1	G546J1_ARATH	G546J1 arabidopsis
94	94.1	566	1	G5PI_SPIOL	G5PI spinacia ol
95	94.1	567	2	O23904_9LILI	O23904 dioscorea t
96	94.1	567	2	O23905_9LILI	O23905 dioscorea t
97	94.1	567	2	O23906_9LILI	O23906 dioscorea t
98	94.1	567	2	O23907_9LILI	O23907 dioscorea t
99	94.1	567	2	O24604_9LILI	O24604 dioscorea t
100	94.1	567	2	O24615_9LILI	O24615 dioscorea t
101	94.1	567	2	Q7DN41_9LILI	Q7DN41 dioscorea t
102	94.1	568	1	G6PIL_CLAAR	G6PIL clarkia arc
103	94.1	568	1	G6PIL_CLAFR	G6PIL clarkia fra
104	94.1	568	1	G6PIL_CLAMI	G6PIL clarkia mil

105	32	94.1	568	1	G6P11_CLAWI	P54239 ciarkia wil	178	31	91.2	473	2	Q69B88 CAMJE	Q69b88 campylobact
106	32	94.1	568	1	G6P11_CLAXA	P54240 ciarkia xan	179	31	91.2	486	2	Q63KP9 BURPS	Q63kp9 burkholderi
107	32	94.1	568	1	G6P11_OENBE	P54243 oenothera m	180	31	91.2	486	2	Q62CD2 BURMA	Q62cd2 burkholderi
108	32	94.1	568	1	Q9SMJ2_9MYRT	Q9smj2 ciarkia gra	181	31	91.2	489	2	Q93383_CAEEL	Q93383 caenorhabdi
109	32	94.1	569	1	G6P11_CLACO	P54235 ciarkia con	182	31	91.2	554	2	Q62100_CAEEL	Q62100 caenorhabdi
110	32	94.1	569	1	G6P11_CLALE	P34796 ciarkia lew	183	31	91.2	554	2	Q92TT6_RHIME	Q92tt6 rhizobium m
111	32	94.1	569	1	G6P12_CLACO	P54241 ciarkia con	184	31	91.2	563	2	Q60CM9_METCA	Q60cm9 methylococc
112	32	94.1	569	1	G6P12_CLALE	P29333 ciarkia lew	185	31	91.2	577	2	Q9BQ5_9RETR	Q9bq5 simian foam
113	32	94.1	569	1	G6P12_CLAXA	P54242 ciarkia xan	186	31	91.2	600	2	Q7ZUZ1_BRARE	Q7zu1 brachydanio
114	32	94.1	569	2	Q8RW71_9MYRT	Q8rw71 ciarkia mod	187	31	91.2	602	1	PEX5_HUMAN	P50542 homo sapien
115	32	94.1	569	2	Q9SMJ1_9MYRT	Q9smj1 ciarkia gra	188	31	91.2	602	2	Q91YC7_MOUSE	Q91yc7 mus musculu
116	32	94.1	569	2	Q8VXR1_CLAEP	Q8vxr1 ciarkia epi	189	31	91.2	603	2	Q92ON3_CRIGR	Q92on3 cricetus
117	32	94.1	569	2	Q8VXR2_CLAEP	Q8vxr2 ciarkia epi	190	31	91.2	608	2	Q642A7_RAT	Q642a7 rattus norv
118	32	94.1	570	1	G6P11_CLARO	P54238 ciarkia ros	191	31	91.2	615	2	Q4JUN3_CORJK	Q4jun3 corynebacte
119	32	94.1	614	2	Q5Z3N5_NOCFA	Q5z3n5 nocardia fa	192	31	91.2	631	2	Q96FN7_HUMAN	Q96fn7 homo sapien
120	32	94.1	834	2	Q8DR73_STRMU	Q8dr73 streptococc	193	31	91.2	633	2	Q55SW1_CRYNE	Q55sw1 cryptococcu
121	32	94.1	834	2	Q4QGF6_LEIMA	Q4qgf6 leishmania	194	31	91.2	633	2	Q7TNJ1_CRIGR	Q7tnj1 cricetus
122	32	94.1	971	2	Q8ITL2_9TRYP	Q8itl2 trypanosoma	195	31	91.2	635	2	Q5KH95_CRYNE	Q5kh95 cryptococcu
123	32	94.1	1142	2	Q57WJ8_9TRYP	Q57wj8 trypanosoma	196	31	91.2	635	2	Q30348_RALSO	Q30348 talstonia s
124	32	94.1	1142	2	Q5BUI3_9TRYP	Q5bui3 trypanosoma	197	31	91.2	639	1	PEX5_MOUSE	Q09012 mus musculu
125	31	91.2	70	2	Q4HDK1_CAMCO	Q4hdk1 campylobact	198	31	91.2	639	2	Q8K3V5_MOUSE	Q8k2v5 mus musculu
126	31	91.2	85	2	Q9RS35_DEIRA	Q9rs35 deinococcus	199	31	91.2	640	2	Q70525_CAVPO	Q70525 cavia porce
127	31	91.2	99	2	Q8FRD0_COREF	Q8frd0 corynebacte	200	31	91.2	640	2	Q92ON5_CRIGR	Q92on5 cricetus
128	31	91.2	105	1	YAS3_SCHPO	Q10139 schizosacch	201	31	91.2	645	2	Q5ZMQ9_CHICK	Q5zmq9 gallus gall
129	31	91.2	129	2	Q8DHX7_SYNEL	Q8dhx7 synechococc	202	31	91.2	655	1	ACSA_NITEU	Q82835 nitrosomona
130	31	91.2	158	1	YBJN_ECOLI	P75815 escherichia	203	31	91.2	663	2	Q81VJ8_HUMAN	Q81v18 homo sapien
131	31	91.2	158	2	Q57R86_SALCH	Q57r86 salmonella	204	31	91.2	669	2	Q8DKV3_SYNEL	Q8dkv3 synechococc
132	31	91.2	158	2	Q5PGM6_SALPA	Q5pgm6 salmonella	205	31	91.2	677	2	Q4T3E0_TETNG	Q4t3e0 tetraodon n
133	31	91.2	158	2	Q8Z847_SALTI	Q8z847 salmonella	206	31	91.2	717	2	Q6FVE9_CANGA	Q6fve9 candida gla
134	31	91.2	158	2	Q8ZQK1_SALTY	Q8zqk1 salmonella	207	31	91.2	721	2	Q6BZ26_YARLI	Q6bz26 yarrowia li
135	31	91.2	179	2	Q90VP2_PLEPL	Q90vf2 pleuronecte	208	31	91.2	738	2	Q4SNF4_TETNG	Q4snf4 tetraodon n
136	31	91.2	208	2	Q71T82_BPPI	Q71t82 bacterioph	209	31	91.2	744	2	Q5MUJ4_LEGPA	Q5mu14 legionella
137	31	91.2	230	2	Q500B3_PSESY	Q500b3 pseudomonas	210	31	91.2	744	2	Q5X334_LEGPA	Q5x334 legionella
138	31	91.2	234	2	Q4K3V9_PSEFS	Q4k3v9 pseudomonas	211	31	91.2	776	2	Q5ZTB4_LEGPH	Q5ztb4 legionella
139	31	91.2	235	2	Q88VX4_LACPL	Q88vx4 lactobacill	212	31	91.2	776	2	Q6K683_ORYSA	Q6k683 oryza sativ
140	31	91.2	241	2	Q5D9T2_SCHJA	Q5d9t2 schistosoma	213	31	91.2	848	2	Q4RVV3_TETNG	Q4rvv3 tetraodon n
141	31	91.2	257	1	Y1985_XYLFT	Q8746 xyella fas	214	31	91.2	863	2	Q9ST50_MAIZE	Q9st50 zea mays (m
142	31	91.2	257	1	Y2609_XYLFA	Q9paas xyella fas	215	31	91.2	865	2	Q8HOG9_TOBAC	Q8hog9 nicotiana t
143	31	91.2	265	2	Q894D1_CLOTE	Q894d1 clostridium	216	31	91.2	878	2	Q4SNF6_TETNG	Q4snf6 tetraodon n
144	31	91.2	287	2	Q9KBZ5_BACHD	Q9kbz5 bacillus ha	217	31	91.2	936	2	Q9LPD9_ARATH	Q9lpd9 arabidopsis
145	31	91.2	291	2	Q611B8_CAEHR	Q611b8 caenorhabdi	218	31	91.2	944	2	Q5FQ83_GLUOX	Q5fq83 gluconobact
146	31	91.2	291	2	Q7M7W7_WOLSU	Q7m7w7 wolinnella s	219	31	91.2	955	2	Q6QNH2_WHEAT	Q6qnh2 triticum ae
147	31	91.2	292	2	Q5MPQ0_SHEEP	Q5mpq0 ovis aries	220	31	91.2	965	2	Q7Q313_ANOGA	Q7q313 anopheles g
148	31	91.2	304	2	Q5LT89_SILPO	Q5lt89 silicibacte	221	31	91.2	968	2	Q6GSA6_BAHE	Q6gsa6 bartonella
149	31	91.2	308	1	XERC_CORGL	Q8nzn9 corynebacte	222	31	91.2	980	2	Q83X74_VIBAN	Q83x74 vibrio angu
150	31	91.2	309	2	Q8WV98_HUMAN	Q8wv98 homo sapien	223	31	91.2	984	2	Q9LZV1_ARATH	Q9lzv1 arabidopsis
151	31	91.2	311	2	Q5NYT6_AZOSE	Q5nyt6 azocarcus sp	224	31	91.2	1094	1	EMBA_MYCTO	P0a561 mycobacteri
152	31	91.2	312	2	Q978P0_TETNG	Q978p0 thermoplasm	225	31	91.2	1094	1	EMBA_MYCTO	P0a560 mycobacteri
153	31	91.2	345	2	Q4T3D9_TETNG	Q4t3d9 tetraodon n	226	31	91.2	1154	2	Q4QOD3_SCHWA	Q4qgd3 schistosoma
154	31	91.2	359	2	Q5V5G5_HALMA	Q5v5g5 haloarcula	227	31	91.2	1230	2	Q889H2_PSESM	Q889h2 pseudomonas
155	31	91.2	362	2	Q99K17_MOUSE	Q99k17 mus musculu	228	31	91.2	1347	2	Q4L263_9BURK	Q4l263 burkholderi
156	31	91.2	365	2	Q7MRK7_WOLSU	Q7mrk7 wolinnella s	229	31	91.2	1391	2	Q4N6K3_THEPA	Q4n6k3 theileria p
157	31	91.2	368	2	Q88Z21_LACPL	Q88z21 lactobacill	230	31	91.2	1413	2	Q4UW10_THEAN	Q4uj10 theileria a
158	31	91.2	384	1	NIFS_RHOCA	Q07177 rhodobacter	231	31	91.2	1447	2	Q4FWA4_LEIMA	Q4fw44 leishmania
159	31	91.2	391	2	Q4LLF8_9BURK	Q4llf8 burkholderi	232	31	91.2	1600	2	Q4XWP5_PLACH	Q4xwp5 plasmodium
160	31	91.2	406	2	Q5KMO0_CRYNE	Q5kmo0 cryptococcu	233	31	91.2	1633	2	Q51NH6_MAGGR	Q51nh6 magnaporthe
161	31	91.2	406	2	Q5KAV1_CRYNE	Q5kav1 cryptococcu	234	31	91.2	1633	2	Q9Y4D8_HUMAN	Q9y4d8 homo sapien
162	31	91.2	409	1	TGFB3_PIG	P15203 sus scrofa	235	31	91.2	3006	2	Q5DIS7_PSEAE	Q5dis7 pseudomonas
163	31	91.2	410	1	TGFB3_MOUSE	P17125 mus musculu	236	31	91.2	3680	2	Q5DIU0_PSEAE	Q5diu0 pseudomonas
164	31	91.2	410	2	Q7SZV3_BRARE	Q7szv3 brachydanio	237	31	91.2	3680	2	Q91157_PSEAE	Q91157 pseudomonas
165	31	91.2	410	2	Q66123_BRACH	Q66123 brachydanio	238	31	91.2	4342	2	Q4Z329_PLAEB	Q4z329 plasmodium
166	31	91.2	412	1	TGFB3_CHICK	P16047 gallus gall	239	31	91.2	4889	2	Q4Z329_PLAEB	Q4z329 plasmodium
167	31	91.2	412	1	TGFB3_HUMAN	P10600 homo sapien	240	31	91.2	4971	2	Q81BG1_PLAF7	Q81bg1 plasmodium
168	31	91.2	412	1	TGFB3_RAT	P10600 homo sapien	241	31	91.2	5054	2	Q7RRU1_PLAYO	Q7rru1 plasmodium
169	31	91.2	412	2	Q91YU7_MOUSE	Q91yu7 mus musculu	242	31	91.2	88.2	45	Q91QU3_9GAMA	Q91qu3 human herpe
170	31	91.2	412	2	Q56A31_RAT	Q56a31 rattus norv	243	30	88.2	88.2	45	Q91QU5_9GAMA	Q91qu5 human herpe
171	31	91.2	414	1	CPXA_PSEPU	Q96a31 rattus norv	244	30	88.2	88.2	46	Q91Q80_9GAMA	Q91q80 pongine her
172	31	91.2	426	2	Q8X013_NEUCR	P00183 pseudomonas	245	30	88.2	81	2	Q6U9H9_9CAUD	Q6u9h9 bacterioph
173	31	91.2	429	2	Q582H7_9TRYP	Q8x013 neurospora	246	30	88.2	81	2	Q56EP0_9CAUD	Q56ep0 aeromonas p
174	31	91.2	440	2	Q8DHL6_SYNEL	Q8dhl6 trypanosoma	247	30	88.2	91	2	Q7N8L3_PHOLL	Q7n8l3 photorhabdu
175	31	91.2	442	2	Q4SXV3_TETNG	Q4sxv3 synechococc	248	30	88.2	105	2	Q4LBU6_PSESH	Q4lbj6 pseudomonas
176	31	91.2	444	2	Q4RLV8_TETNG	Q4rlv8 tetraodon n	249	30	88.2	105	2	Q4ZMF3_PSESH	Q4zmf3 pseudomonas
177	31	91.2	473	2	Q69BD7_CAMCO	Q69bd7 campylobact	250	30	88.2	110	2	Q69141_9GAMA	Q69141 human herpe
										120	1	E413_AD812	P36709 human adeno

251	30	88.2	136	2	Q9DAL0_MOUSE	Q9dal0 mus musculus	324	30	88.2	346	2	Q8X8Z2_ECO57	Q8x8z2 escherichia
252	30	88.2	139	2	Q5TP24_ANOGA	Q5tp24 anopheles g	325	30	88.2	350	2	Q6PJS6_HUMAN	Q6pjs6 homo sapien
253	30	88.2	156	2	Q7NWU0_CHRVO	Q7nwu0 chromobacte	326	30	88.2	353	2	Q6XN51_RHOER	Q6xn51 rhodococcus
254	30	88.2	158	2	Q9X0A5_THEMEA	Q9x0a5 thermotoga	327	30	88.2	357	2	Q4R6X6_MACFA	Q4r6x6 macaca fasc
255	30	88.2	166	2	Q5DFY8_SCHJA	Q5dfy8 schistosoma	328	30	88.2	364	2	Q8R9W51_BRAJA	Q8r9w51 bradyrhizob
256	30	88.2	169	2	Q76LM7_HYDTN	Q76lm7 hydropenoba	329	30	88.2	366	2	Q8XXM1_RALSO	Q8xxm1 ralstonia s
257	30	88.2	176	2	Q5DA00_SCHJA	Q5da00 schistosoma	330	30	88.2	370	2	Q8XND3_RHOPA	Q8xnd3 rhodopseudo
258	30	88.2	178	2	Q84XL1_LVCES	Q84xl1 lycopersico	331	30	88.2	372	2	Q6ZUR8_PYRAE	Q6zur8 pyrobaculum
259	30	88.2	181	2	Q6GCD0_STAAS	Q6gcd0 staphylococ	332	30	88.2	373	2	Q8C770_MOUSE	Q8c770 mus musculus
260	30	88.2	181	2	Q6GJX6_STAAR	Q6gjx6 staphylococ	333	30	88.2	376	2	Q9C106_MOUSE	Q9c106 mus musculus
261	30	88.2	181	2	Q5HIV2_STAAC	Q5hiv2 staphylococ	334	30	88.2	376	2	Q6QIB7_BURKHO	Q6qib7 burkholderi
262	30	88.2	181	2	Q7A7M4_STAAN	Q7a7m4 staphylococ	335	30	88.2	376	2	Q4NFC4_9MICC	Q4nfc4 athrobacte
263	30	88.2	181	2	Q8NYA4_STAAN	Q8nya4 staphylococ	336	30	88.2	376	2	Q8UAR2_AGR75	Q8uar2 agrobacteri
264	30	88.2	181	2	Q99WN5_STAAM	Q99wn5 staphylococ	337	30	88.2	376	2	Q4FXP7_LEIMA	Q4fxp7 leishmania
265	30	88.2	181	2	Q5XHD0_XENLA	Q5xhd0 xenopus lae	338	30	88.2	376	2	Q5X4W2_LEGPA	Q5x4w2 legionella
266	30	88.2	181	2	Q9VJ70_DROME	Q9vj70 drosophila	339	30	88.2	376	2	Q5X4W2_LEGPA	Q5x4w2 legionella
267	30	88.2	192	2	Q5RHX5_BRARE	Q5rhx5 brachydanio	340	30	88.2	376	2	Q4R6H4_MACFA	Q4r6h4 macaca fasc
268	30	88.2	196	2	Q7Z6B8_HUMAN	Q7z6b8 homo sapien	341	30	88.2	376	2	Q5YXZ5_NOCPA	Q5yxz5 nocardia fa
269	30	88.2	197	2	Q5NYK1_AZOSE	Q5nyk1 azoarcus sp	342	30	88.2	376	2	Q8WUF8_HUMAN	Q8wuf8 homo sapien
270	30	88.2	201	2	Q5NYK1_AZOSE	Q5nyk1 azoarcus sp	343	30	88.2	376	2	Q5ZK44_CHICK	Q5zk44 gallus gall
271	30	88.2	208	2	Q6PAM8_MOUSE	Q6pam8 mus musculus	344	30	88.2	376	2	Q82BK6_STRAW	Q82bk6 streptomyce
272	30	88.2	210	2	Q6PTI9_9TURB	Q6pti9 stylochoas s	345	30	88.2	376	2	Q4SKG9_TETNG	Q4skg9 tetradon n
273	30	88.2	210	2	Q664A1_YERPS	Q664a1 yersinia ps	346	30	88.2	376	2	Q6PFX1_MOUSE	Q6pfx1 mus musculus
274	30	88.2	224	2	Q4IF53_GIBZE	Q4if53 gibberella	347	30	88.2	376	2	Q9HUL6_PSEAE	Q9hul6 pseudomonas
275	30	88.2	229	2	Q45898_COCBU	Q45898 coxiella bu	348	30	88.2	376	2	Q9RD01_STRCO	Q9rd01 streptomyce
276	30	88.2	231	2	Q28500_MACBU	Q28500 macaca mula	349	30	88.2	376	2	Q92T96_RHIME	Q92t96 rhizobium m
277	30	88.2	233	2	Q52877_COXBU	Q52877 coxiella bu	350	30	88.2	376	2	Q8DS32_STRMU	Q8ds32 streptococc
278	30	88.2	233	2	Q45912_COXBU	Q45912 coxiella bu	351	30	88.2	376	2	Q6FJMA_CANGA	Q6fjma candida gla
279	30	88.2	233	2	Q45935_COXBU	Q45935 coxiella bu	352	30	88.2	376	2	Q9YH86_CHICK	Q9yh86 gallus gall
280	30	88.2	233	2	Q67SJ6_SYMTN	Q67sj6 symbiobacte	353	30	88.2	376	2	Q53J22_ORYSA	Q53j22 oryza sativ
281	30	88.2	236	2	Q83A05_COXBU	Q83a05 coxiella bu	354	30	88.2	376	2	Q5FPU0_GLUOC	Q5fpu0 gluconobact
282	30	88.2	239	2	Q26768_METTH	Q26768 methanobact	355	30	88.2	376	2	Q8XIK0_COPCI	Q8xik0 coprinus ci
283	30	88.2	252	2	Q8A4J0_BACTN	Q8a4j0 bacteroides	356	30	88.2	376	2	Q5P3I6_AZOSE	Q5p3i6 azoarcus sp
284	30	88.2	261	2	Q59147_9GAMA	Q59147 human herpe	357	30	88.2	376	2	Q6ATU5_ORYSA	Q6at5 oryza sativ
285	30	88.2	263	2	Q60407_CRICR	Q60407 cricetus cr	358	30	88.2	376	2	Q44317_9ARAC	Q44317 dugesiella
286	30	88.2	272	2	Q82929_STRAW	Q82929 streptomyce	359	30	88.2	376	2	Q8R9N8_THETN	Q8r9n8 thermoanaer
287	30	88.2	276	2	Q6UXP7_HUMAN	Q6uxp7 homo sapien	360	30	88.2	376	2	Q6FXL9_CANGA	Q6fxl9 candida gla
288	30	88.2	281	2	Q76D05_STRGR	Q76d05 streptomyce	361	30	88.2	376	2	Q69125_9GAMA	Q69125 human herpe
289	30	88.2	284	2	Q5CE87_CRYHO	Q5ce87 cryptospori	362	30	88.2	376	2	Q8AZK7_9GAMA	Q8azk7 mycobacteri
290	30	88.2	284	2	Q9HUX6_PSEAE	Q9hux6 pseudomonas	363	30	88.2	376	2	Q53470_MYCTU	Q53470 mycobacteri
291	30	88.2	289	2	Q8LBE21_ARATH	Q8lbe21 arabidopsis	364	30	88.2	376	2	Q5MAM6_PARUM	Q5mam6 paracalamyd
292	30	88.2	291	2	Q9GR92_ARTSF	Q9gr92 artemia san	365	30	88.2	376	2	Q5WAF7_BACSK	Q5waf7 bacillus cl
293	30	88.2	292	2	Q9SA78_ARATH	Q9sa78 arabidopsis	366	30	88.2	376	2	PGS1_YEAST	PGS1 saccharomyc
294	30	88.2	292	2	Q8CCA0_MOUSE	Q8cca0 mus musculus	367	30	88.2	376	2	Q82507_ARATH	Q82507 arabisidopsi
295	30	88.2	301	2	Q4K8S1_PSEF5	Q4k8s1 pseudomonas	368	30	88.2	376	2	Q82190_STRAW	Q82190 streptomyce
296	30	88.2	301	2	Q9CMH0_MOUSE	Q9cmh0 mus musculus	369	30	88.2	376	2	Q4P1M0_USTMA	Q4p1m0 ustilago ma
297	30	88.2	305	2	Q5RHX6_BRARE	Q5rhx6 brachydanio	370	30	88.2	376	2	Q5TVV8_ANOGA	Q5tvv8 anopheles g
298	30	88.2	306	2	Q4ZUQ4_PSESY	Q4zuq4 pseudomonas	371	30	88.2	376	2	Q8C914_MOUSE	Q8c914 mus musculus
299	30	88.2	306	2	Q8EJ00_SHEON	Q8ej00 shewanella	372	30	88.2	376	2	Q65UK0_MANSM	Q65uk0 manheimia
300	30	88.2	306	2	Q883S3_PSESM	Q883s3 pseudomonas	373	30	88.2	376	2	Q6XPA6_PHYIN	Q6xpa6 phytophthor
301	30	88.2	306	2	Q8C5X2_MOUSE	Q8c5x2 mus musculus	374	30	88.2	376	2	Q6XPA7_PHYIN	Q6xpa7 phytophthor
302	30	88.2	308	2	Q69136_9GAMA	Q69136 human herpe	375	30	88.2	376	2	Q6XPA9_PHYIN	Q6xpa9 phytophthor
303	30	88.2	309	2	Q92564_HUMAN	Q92564 homo sapien	376	30	88.2	376	2	Q6XPB0_9STRA	Q6xpb0 phytophthor
304	30	88.2	315	2	Q723F3_HUMAN	Q723f3 homo sapien	377	30	88.2	376	2	Q6XPB1_9STRA	Q6xpb1 phytophthor
305	30	88.2	316	2	Q5FPU4_GLIJX	Q5fpu4 gluconobact	378	30	88.2	376	2	Q6XPB2_9STRA	Q6xpb2 phytophthor
306	30	88.2	318	2	Q5BEU1_EMENT	Q5beu1 aspergillus	379	30	88.2	376	2	Q6XPB3_9STRA	Q6xpb3 phytophthor
307	30	88.2	318	2	Q5H033_XANOR	Q5h033 xanthomonas	380	30	88.2	376	2	Q6XPB4_PHYIN	Q6xpb4 phytophthor
308	30	88.2	318	2	Q87D99_XYLFT	Q87d99 xyella fas	381	30	88.2	376	2	Q6XPB5_PHYIN	Q6xpb5 phytophthor
309	30	88.2	318	2	Q8PM38_XANAC	Q8pm38 xanthomonas	382	30	88.2	376	2	Q6XPB6_PHYIN	Q6xpb6 phytophthor
310	30	88.2	319	2	Q9PB01_XYLFA	Q9pb01 xyella fas	383	30	88.2	376	2	Q6XPB7_PHYIN	Q6xpb7 phytophthor
311	30	88.2	321	2	Q4UT87_XANCP	Q4ut87 xanthomonas	384	30	88.2	376	2	Q6XPB8_PHYIN	Q6xpb8 phytophthor
312	30	88.2	321	2	Q8PAD8_XANCP	Q8pad8 xanthomonas	385	30	88.2	376	2	Q6XPB9_PHYIN	Q6xpb9 phytophthor
313	30	88.2	323	1	LRP16_MOUSE	Q92xb1 mus musculus	386	30	88.2	376	2	Q6XPC0_PHYIN	Q6xpc0 phytophthor
314	30	88.2	326	2	Q7WXM1_ALCEU	Q7wxm1 alcaligenes	387	30	88.2	376	2	Q6XPC1_PHYIN	Q6xpc1 phytophthor
315	30	88.2	326	2	Q4LR14_9BURK	Q4lr14 burkholderi	388	30	88.2	376	2	Q6XPC3_9STRA	Q6xpc3 phytophthor
316	30	88.2	327	2	Q5LR05_SILPO	Q5lr05 silicibacte	389	30	88.2	376	2	Q6XPC5_PHYIN	Q6xpc5 phytophthor
317	30	88.2	329	2	Q7SB66_NEUCR	Q7sb66 neurospora	390	30	88.2	376	2	Q6XPC6_PHYIN	Q6xpc6 phytophthor
318	30	88.2	332	1	EXPA_CEPAC	P11935 c cephalosp	391	30	88.2	376	2	Q6XPC7_PHYIN	Q6xpc7 phytophthor
319	30	88.2	332	2	Q9P4T5_CEPAC	Q9p4t5 cephalosp	392	30	88.2	376	2	Q8ANE6_PHYIN	Q8ane6 phytophthor
320	30	88.2	334	2	Q5V1K1_HALMA	Q5v1k1 haloarcula	393	30	88.2	376	2	Q4PGB7_USTMA	Q4pgb7 ustilago ma
321	30	88.2	338	2	Q9EYD4_ECO57	Q9eyd4 escherichia	394	30	88.2	376	2	Q6BHC9_SOLITU	Q6bhc9 solanum tub
322	30	88.2	338	2	Q4LPN5_9BURK	Q4lpn5 burkholderi	395	30	88.2	376	2	Q6P3G2_BRARE	Q6p3g2 brachydanio
323	30	88.2	342	2	Q5YMK5_NOCPA	Q5ymk5 nocardia fa	396	30	88.2	376	2	Q86DC0_CABEL	Q86dc0 caenorhabdi

397	30	88.2	595	2	Q71683_9CAUD	Q71683 mycobacteri	470	30	88.2	1900	2	Q8WTD2_DROME	Q8WTD2 drosophila
398	30	88.2	601	2	Q5CTA2_CRYPV	Q5cta2 cryptospori	471	30	88.2	1900	2	Q9VTS9_DROME	Q9VTS9 drosophila
399	30	88.2	601	2	Q6D0G9_ERWCT	Q6d0g9 erwinia car	472	30	88.2	1915	2	Q5VAF6_HUMAN	Q5VAF6 homo sapien
400	30	88.2	602	2	Q6CF71_YARLI	Q6cf71 yarrowia li	473	30	88.2	1947	2	Q4SIP4_TETNG	Q4SIP4 tetradodon n
401	30	88.2	608	2	Q8QZW3_MOUSE	Q8qzw3 mus musculu	474	30	88.2	2072	2	Q4SAO9_TETNG	Q4SAO9 tetradodon n
402	30	88.2	613	1	GLSA1_BRAJA	Q8qna7 bradyrhizob	475	30	88.2	2145	2	Q4RHK4_TETNG	Q4RHK4 tetradodon n
403	30	88.2	614	2	Q69HN8_CIOIN	Q69hn8 ciona intes	476	30	88.2	2179	2	Q91DM0_9VIRU	Q91DM0 petunia vei
404	30	88.2	614	2	Q6D6M5_ERWCT	Q6d6m5 erwinia car	477	30	88.2	2180	2	Q6XKE6_9VIRU	Q6XKE6 petunia vei
405	30	88.2	619	2	Q73811_MYCPA	Q73811 mycobacteri	478	30	88.2	2202	2	Q59EQ7_HUMAN	Q59EQ7 homo sapien
406	30	88.2	628	2	Q5JM56_ORYSA	Q5jm56 oryza sativ	479	30	88.2	2592	2	Q84T10_ORYSA	Q84T10 oryza sativ
407	30	88.2	632	2	Q751J9_ORYSA	Q751j9 oryza sativ	480	30	88.2	3038	1	TRIO_HUMAN	Q75962 homo sapien
408	30	88.2	632	2	Q57K30_SALCH	Q57K30 salmonella	481	30	88.2	3072	2	Q6LEZ2_PLAF7	Q6LEZ2 plasmodium
409	30	88.2	639	2	Q9Z417_STRSQ	Q9z417 streptomyce	482	30	88.2	3470	2	Q88F79_PSEPK	Q88F79 pseudomonas
410	30	88.2	639	2	Q8KNM2_MOUSE	Q8knm2 mus musculu	483	30	88.2	4034	2	Q7R5J5_GIALA	Q7R5J5 giardia lam
411	30	88.2	656	2	Q4Q0K8_LEIMA	Q4q0k8 leishmania	484	30	88.2	6260	2	Q54299_STRHY	Q54299 streptomyce
412	30	88.2	658	1	SPEA_ECO57	Q8qcx9 escherichia	485	30	88.2	8348	2	Q4S6V5_TETNG	Q4S6V5 tetradodon n
413	30	88.2	658	1	SPEA_ECOL6	Q8fe34 escherichia	486	30	88.2	8563	2	Q54297_STRHY	Q54297 streptomyce
414	30	88.2	658	1	SPEA_ECOLI	P21170 escherichia	487	29	85.3	42	Q8EJZ9_SHEON	Q8EJZ9 shewanella	
415	30	88.2	658	1	SPEA_SALTY	P60658 salmonella	488	29	85.3	67	Q51A41_ENTHI	Q51A41 entamoeba h	
416	30	88.2	658	1	SPEA_SALTY	P60659 salmonella	489	29	85.3	74	Q7VK46_HELHP	Q7VK46 helicobacte	
417	30	88.2	659	1	SPEA_YERPE	Q8zhg8 yersinia pe	490	29	85.3	80	Q9KRZ6_VIBCH	Q9KRZ6 vibrio chol	
418	30	88.2	661	2	Q6AX23_XENLA	Q6ax23 xenopus lae	491	29	85.3	81	Q6EB44_CAMJE	Q6EB44 campylobact	
419	30	88.2	662	1	SPEA_SHIFL	Q83q93 shigella fl	492	29	85.3	84	O80132_9CAUD	O80132 bacteriopho	
420	30	88.2	662	2	Q4WU83_ASPFU	Q4wu83 aspergillus	493	29	85.3	85	Q21893_9CAUD	Q21893 bacteriopho	
421	30	88.2	682	2	Q8BK93_MOUSE	Q8bk93 mus musculu	494	29	85.3	85	O8FN33_COREF	O8FN33 corynebacte	
422	30	88.2	706	2	Q8VNN5_9VIRU	Q8vnn5 bacteriopho	495	29	85.3	87	Q9DFJ7_GILMI	Q9DFJ7 gillichthys	
423	30	88.2	712	2	Q6C6Z8_YARLI	Q6c6z8 yarrowia li	496	29	85.3	90	Y4IG_RHISN	Y55490 rhizobium s	
424	30	88.2	713	2	Q18662_CABEL	Q18662 caenorhabdi	497	29	85.3	94	Q63NA3_BURPS	Q63na3 burkholderi	
425	30	88.2	733	2	Q5XQP4_MAIZE	Q5xqf4 zea mays (m	498	29	85.3	95	Q86SR9_HUMAN	Q86sr9 homo sapien	
426	30	88.2	736	2	Q8KNK0_SALTI	Q8knk0 salmonella	499	29	85.3	99	Q7NJ12_GLOVI	Q7NJ12 gloeobacter	
427	30	88.2	745	2	Q4S3N3_TETNG	Q4s3n3 tetradodon n	500	29	85.3	99	Q7NMC1_GLOVI	Q7NMC1 gloeobacter	
428	30	88.2	751	2	Q6BXA9_DEBHA	Q6bx93 debaryomyce	501	29	85.3	102	Q4KA72_PSEFS	Q4KA72 pseudomonas	
429	30	88.2	756	2	Q4IZG3_AZOVI	Q4izg3 azotobacter	502	29	85.3	102	Q9J9C0_9LUTE	Q9J9C0 sugarcane y	
430	30	88.2	774	2	Q7QC99_ANOGA	Q7qcc9 anopheles g	503	29	85.3	106	Y04K_BP74	Y07069 bacteriopho	
431	30	88.2	781	2	Q5AZZ3_EMENI	Q5azz3 aspergillus	504	29	85.3	109	Q8KUM1_SYN7	Q8KUM1 synecchococ	
432	30	88.2	830	2	Q5YRU2_NOCFA	Q5yru2 nocardia fa	505	29	85.3	114	Q7BF85_YERPE	Q7BF85 yersinia ps	
433	30	88.2	897	2	Q7U1L0_MYCBO	Q7u1l0 mycobacteri	506	29	85.3	114	Q93PD5_YERPE	Q93PD5 yersinia ps	
434	30	88.2	912	2	Q8CCJ2_MOUSE	Q8ccj2 mus musculu	507	29	85.3	114	Q66A07_YERPS	Q66A07 yersinia ps	
435	30	88.2	923	2	Q5E1V0_VIBF1	Q5e1v0 vibrio fisc	508	29	85.3	115	Q8SOG0_ORYSA	Q8SOG0 oryza sativ	
436	30	88.2	952	2	Q9LQ93_ARATH	Q9lq93 arabidopsis	509	29	85.3	116	Q7P7E0_FUSNV	Q7P7E0 fusobacteri	
437	30	88.2	975	2	Q4WNK0_ASPFU	Q4wnk0 aspergillus	510	29	85.3	120	Q8CL76_YERPE	Q8CL76 yersinia pe	
438	30	88.2	1009	2	Q7SEV7_NEUCR	Q7sev7 neurospora	511	29	85.3	121	Q69SQ6_9ADEN	Q69SQ6 simian aden	
439	30	88.2	1026	2	Q86T76_HUMAN	Q86t76 homo sapien	512	29	85.3	121	Q6QP86_9ADEN	Q6QP86 simian aden	
440	30	88.2	1027	1	DOCK1_MOUSE	Q8bur4 mus musculu	513	29	85.3	121	Q6QPC2_9ADEN	Q6QPC2 simian aden	
441	30	88.2	1033	2	Q4JZV5_AZOVI	Q4jzv5 azotobacter	514	29	85.3	121	Q6QPF8_9ADEN	Q6QPF8 simian aden	
442	30	88.2	1040	2	Q9T014_ARATH	Q9t014 arabidopsis	515	29	85.3	121	Q8UY65_9ADEN	Q8UY65 simian aden	
443	30	88.2	1083	2	Q9RPH6_MYCSM	Q9rph6 mycobacteri	516	29	85.3	122	Q6HIA5_9ADEN	Q6HIA5 human aden	
444	30	88.2	1094	1	EX5B_MYCTU	P96920 mycobacteri	517	29	85.3	122	Q5VH99_ADE04	Q5VH99 human aden	
445	30	88.2	1099	2	Q571B9_MOUSE	Q571b9 mus musculu	518	29	85.3	124	Q8U618_AGR75	Q8U618 agrobacteri	
446	30	88.2	1108	2	Q4SEAT_TETNG	Q4sear tetradodon n	519	29	85.3	126	Q7QYD0_GIALA	Q7QYD0 giardia lam	
447	30	88.2	1123	2	Q43948_PLAFA	Q43948 plasmodium	520	29	85.3	131	Q937P9_SALET	Q937P9 salmonella	
448	30	88.2	1189	2	Q7URB6_RHOBA	Q7urb6 rhodopirell	521	29	85.3	131	Q70DP1_ACTPL	Q70DP1 actinobacil	
449	30	88.2	1206	2	Q4IM26_GIBBEZ	Q4im26 gibberella	522	29	85.3	131	Q70DP2_ACTLI	Q70DP2 actinobacil	
450	30	88.2	1294	2	Q4RFA4_TETNG	Q4rfa4 tetradodon n	523	29	85.3	133	Q4ZPV7_PSESY	Q4ZPV7 pseudomonas	
451	30	88.2	1334	2	Q4RTS6_TETNG	Q4rts6 tetradodon n	524	29	85.3	134	Q88615_PSESM	Q88615 pseudomonas	
452	30	88.2	1364	2	Q8PUN8_XANAC	Q8pjn8 xanthomonas	525	29	85.3	137	Q50LS9_ENTHI	Q50LS9 entamoeba h	
453	30	88.2	1402	2	Q5GZ20_XANOR	Q5gzz20 xanthomona	526	29	85.3	143	Q59287_PYRHO	Q59287 pyrococcus	
454	30	88.2	1424	2	Q6P9K6_MOUSE	Q6p9k6 mus musculu	527	29	85.3	145	Q5VP06_ORYSA	Q5VP06 oryza sativ	
455	30	88.2	1566	2	Q8MZD0_DROME	Q8mzd0 drosophila	528	29	85.3	145	Q5P0R6_AZOSE	Q5P0R6 azoarcus sp	
456	30	88.2	1566	2	Q9VJT4_DROME	Q9vjt4 drosophila	529	29	85.3	149	Q9Z2R2_MOUSE	Q9Z2R2 mus musculu	
457	30	88.2	1569	2	Q8WTD1_DROME	Q8vtd1 drosophila	530	29	85.3	155	Q4WLQ8_ASPFU	Q4WLQ8 aspergillus	
458	30	88.2	1569	2	Q9VTS8_DROME	Q9vts8 drosophila	531	29	85.3	155	Q834C9_ENTFA	Q834C9 enterococu	
459	30	88.2	1577	2	Q9NKC7_DROME	Q9nkc7 drosophila	532	29	85.3	156	O8LX8_ARATH	O8LX8 arabidopsis	
460	30	88.2	1606	2	Q9XB15_MYCBO	Q9xb15 mycobacteri	533	29	85.3	158	Q6U9T5_9CAUD	Q6U9T5 bacteriopho	
461	30	88.2	1606	2	Q7D7L9_MYCTU	Q7d7l9 mycobacteri	534	29	85.3	158	Q56E25_9CAUD	Q56E25 aeromonas p	
462	30	88.2	1606	2	Q7TZ47_MYCBO	Q7tz47 mycobacteri	535	29	85.3	159	O8SS75_ENCCU	O8SS75 encephalito	
463	30	88.2	1653	2	Q5YM84_NOCFA	Q5ym84 nocardia fa	536	29	85.3	159	Q67TG9_SYMTH	Q67TG9 symbiobacte	
464	30	88.2	1669	2	Q7N7D7_PHOLL	Q7n7d7 photorhabdu	537	29	85.3	161	O5B6X4_EMENI	O5B6X4 aspergillus	
465	30	88.2	1670	2	Q8WS23_DROME	Q8ws23 drosophila	538	29	85.3	162	Q7N9X2_PHOLL	Q7N9X2 photorhabdu	
466	30	88.2	1670	2	Q9VTS6_DROME	Q9vts6 drosophila	539	29	85.3	163	Q873A6_NEUCK	Q873A6 neuropsora	
467	30	88.2	1865	1	DOCK1_HUMAN	Q14185 homo sapien	540	29	85.3	165	Q66637_AQUAE	Q66637 aquifex aeo	
468	30	88.2	1886	2	Q6XXE5_9VIRU	Q6xxe5 petunia vei	541	29	85.3	169	Q41938_GIBBEZ	Q41938 gibberella	
469	30	88.2	1900	2	Q8WS24_DROME	Q8ws24 drosophila	542	29	85.3	170	O51MR3_MAGGR	O51MR3 magnaporthe	

543	29	85.3	170	2	Q7N6G7	PHOLL	Q7n6g7	photorhabdu	616	29	85.3	244	2	Q7WP46	BORBR	Q7wp46	bordetella
544	29	85.3	170	2	Q7NS63	CHRV0	Q7ns63	chromobacte	617	29	85.3	244	2	Q4V8S1	BRARE	Q4v8s1	brachydanio
545	29	85.3	172	2	Q5E537	VIBF1	Q5e537	vibrio fisc	618	29	85.3	245	2	Q8EPY9	OCEIH	Q8epy9	oceanobacil
546	29	85.3	172	2	Q3K3K2	BURPS	Q3k3k2	burkholderi	619	29	85.3	246	2	Q41VT5	AZOV1	Q41vt5	azotobacter
547	29	85.3	174	2	Q7ND96	GLOV1	Q7nd96	gloeobacter	620	29	85.3	248	2	Q4TKN9	9SPHN	Q4tkn9	erythrobact
548	29	85.3	174	2	Q64D63	MESAU	Q64d63	mesocricetu	621	29	85.3	248	2	Q63VF1	BURPS	Q63vf1	burkholderi
549	29	85.3	175	2	Q4L552	STAHU	Q4l552	staphylococ	622	29	85.3	248	2	Q90XF9	BRARE	Q90xf9	brachydanio
550	29	85.3	175	2	Q7MV46	PORGI	Q7mv46	porphyromon	623	29	85.3	250	2	Q746K8	THER2	Q746k8	thermus the
551	29	85.3	182	2	Q8EK21	OCEIH	Q8ek21	oceanobacil	624	29	85.3	250	2	Q53W72	THER78	Q53w72	thermus the
552	29	85.3	184	2	Q92I16	ECOLI	Q92i16	escherichia	625	29	85.3	250	2	Q413H5	GIBZE	Q413h5	gibberella
553	29	85.3	186	2	Q4NN47	9DEL1	Q4nna7	anaeromyxob	626	29	85.3	251	2	Q9X7N9	STRCO	Q9x7n9	streptomyce
554	29	85.3	189	2	Q61MH5	NEOCA	Q61mh5	neospora ca	627	29	85.3	252	2	Q82PM0	STRAW	Q82pm0	streptomyce
555	29	85.3	192	1	MOBA	NEIMA	Q9jua5	neisseria m	628	29	85.3	252	2	Q8FNW6	CORBP	Q8fnw6	corynebacte
556	29	85.3	192	1	MOBA	NEIMB	P58747	neisseria m	629	29	85.3	252	2	Q77K62	NUCL	Q77k62	helicoverpa
557	29	85.3	192	2	Q7QW59	GIALA	Q7qw59	giardia lam	630	29	85.3	254	2	Q77LK4	9NUCL	Q77lk4	helicoverpa
558	29	85.3	192	2	Q7NT62	CHRV0	Q7nt62	chromobacte	631	29	85.3	254	2	Q99GW8	9NUCL	Q99gw8	helicoverpa
559	29	85.3	193	2	Q7NV13	CHRV0	Q7nv13	chromobacte	632	29	85.3	255	1	Y1632	METJA	Y1632	methanococc
560	29	85.3	196	2	Q5P7P9	AZOSE	Q5p7p9	azoaracus sp	633	29	85.3	255	1	Q88PO7	PSEPK	Q88po7	pseudomonas
561	29	85.3	197	2	Q8PN82	XANAC	Q8pn82	xanthomonas	634	29	85.3	263	2	Q4KEM0	PSEPK	Q4kem0	pseudomonas
562	29	85.3	198	2	Q8XZR6	RALSO	Q8xze6	raistonia s	635	29	85.3	263	2	Q88UJ12	PSEPK	Q88uj12	pseudomonas
563	29	85.3	199	1	TORD	EGOL6	Q8fzr6	escherichia	636	29	85.3	266	2	Q5XGF8	XENTR	Q5xgf8	xenopus tro
564	29	85.3	199	2	Q5N2X9	SYNP6	Q5n2x9	synecochoc	637	29	85.3	266	2	Q6NU97	XENLA	Q6nu97	xenopus lae
565	29	85.3	199	2	Q42037	9GAMA	O42037	alcalaphine	638	29	85.3	269	1	R1BF	MYCPN	R1bf	mycoplasma
566	29	85.3	201	2	Q7CMH7	BACAN	Q7cmh7	bacillus an	639	29	85.3	269	2	O5XHE1	XENLA	O5xhe1	xenopus lae
567	29	85.3	201	2	Q4MKL0	BACCE	Q4mk10	bacillus ce	640	29	85.3	270	2	Q4TAI6	TETNG	Q4tai6	tetradodon n
568	29	85.3	201	2	Q74P26	BACC1	Q74p26	bacillus ce	641	29	85.3	271	2	Q4W477	AERSA	Q4w477	aeromonas s
569	29	85.3	201	2	Q82VG6	NITEU	Q82vg6	nitrosomona	642	29	85.3	271	2	Q699Q3	AERHY	Q699q3	aeromonas h
570	29	85.3	201	2	Q9X317	BACAN	Q9x317	bacillus an	643	29	85.3	271	2	O5YY59	NOCPA	O5yy59	nocardia fa
571	29	85.3	202	2	Q75AD3	ASHGO	Q75ad3	ashbya goss	644	29	85.3	273	2	Q67LW2	SYMTH	Q67lw2	syndiobacte
572	29	85.3	204	2	Q5F8L8	NEIG1	Q5f8l8	neisseria g	645	29	85.3	273	2	Q6P0C0	BRARE	Q6p0c0	brachydanio
573	29	85.3	204	2	Q67KU0	SYMTH	Q67ku0	syndiobacte	646	29	85.3	274	2	Q6NEN3	CORDI	Q6nen3	corynebacte
574	29	85.3	206	2	Q4T545	TETNG	Q4t545	tetradodon n	647	29	85.3	275	2	Q4YWN3	PLABE	Q4ywn3	plasmodium
575	29	85.3	208	2	Q67TGO	SYMTH	Q67tgo	syndiobacte	648	29	85.3	277	2	O58047	PYRHO	O58047	pyrococcus
576	29	85.3	210	2	Q8VP38	ECOLI	Q8vp38	escherichia	649	29	85.3	280	2	Q86LQ0	BRABE	Q86lq0	branchiost
577	29	85.3	210	2	Q76BF5	ORYLA	Q76bf5	oryzias lat	650	29	85.3	280	2	O5N664	SYNP6	O5n664	haemophilus
578	29	85.3	211	2	Q70I46	DIPPU	Q70i46	diploptera	651	29	85.3	283	1	Y16544	HAEIN	Y16544	haemophilus
579	29	85.3	211	2	Q7NL44	GLOV1	Q7nl44	gloeobacter	652	29	85.3	283	2	Q4QJ79	HAEI8	Q4qj79	haemophilus
580	29	85.3	213	2	Q5DDT7	SCHJA	O5ddt7	schistosoma	653	29	85.3	284	2	Q4QPH2	PSEF5	Q4qph2	pseudomonas
581	29	85.3	214	1	Y220B	HAEIN	Q4qmw7	haemophilus	654	29	85.3	285	2	O9HL11	THEAC	O9hl11	thermoplasm
582	29	85.3	214	2	Q4QNW7	HAEI8	Q4qnw7	haemophilus	655	29	85.3	285	2	Q4J2D1	AZOV1	Q4j2d1	azotobacter
583	29	85.3	214	2	Q7NVB7	CHRV0	Q7nvb7	chromobacte	656	29	85.3	285	2	Q9Z867	CHLPN	Q9z867	chlamydia p
584	29	85.3	215	2	Q8NTE3	CORGL	Q8nte3	corynebacte	657	29	85.3	286	2	Q89V19	BRAJA	Q89v19	bradyrhizob
585	29	85.3	216	1	PAAD	STRCO	Q9kyp1	streptomyce	658	29	85.3	287	2	Q4ZVF7	PSESY	Q4zvf7	pseudomonas
586	29	85.3	216	2	Q4WUZ8	ASPFU	Q4wuz8	aspergillus	659	29	85.3	288	1	ZIPA	PSESM	ZIPA	pseudomonas
587	29	85.3	220	2	Q4R883	MACFA	Q4r883	macaca fasc	660	29	85.3	289	2	Q4S363	TETNG	Q4s363	tetradodon n
588	29	85.3	220	2	Q5YUZO	NOCPA	Q5yuz0	nocardia fa	661	29	85.3	290	2	Q5BJN1	RAT	Q5bjn1	rattus norv
589	29	85.3	221	2	Q67M46	SYMTH	Q67m46	syndiobacte	662	29	85.3	291	1	PCTL	HUMAN	PCTL	mus musculu
590	29	85.3	223	2	Q7V5F0	BORPE	Q7v5f0	bordetella	663	29	85.3	291	1	Q4TNL4	9SPHN	Q4tnl4	erythrobact
591	29	85.3	223	2	Q7W5H5	BORPA	Q7w5h5	bordetella	664	29	85.3	295	2	KSGA	STARP	KSGA	s dimethyla
592	29	85.3	223	2	Q7WD11	BORBR	Q7wd11	bordetella	665	29	85.3	296	1	KSGA	STARP	KSGA	theliera p
593	29	85.3	224	2	Q88LM3	PSEPK	Q88lm3	pseudomonas	666	29	85.3	296	2	Q4N4C6	THEPA	Q4n4c6	theileria p
594	29	85.3	226	2	Q62CA9	BURMA	Q62ca9	burkholderi	667	29	85.3	296	2	ZIPA	PSEPK	ZIPA	pseudomonas
595	29	85.3	229	2	Q7P6L7	FUSNV	Q7p6l7	fusobacteri	668	29	85.3	298	2	O5YPT6	NOCPA	O5ypt6	nocardia fa
596	29	85.3	229	2	Q4HBD4	9DEIO	Q4hbd4	deinococcus	669	29	85.3	299	2	O89E74	BRAJA	O89e74	bradyrhizob
597	29	85.3	229	2	Q8REM9	FUSNV	Q8rem9	fusobacteri	670	29	85.3	299	2	O9L526	VIBCH	O9l526	vibrio chol
598	29	85.3	231	2	Q5G189	XANOR	Q5g189	xanthomonas	671	29	85.3	301	2	Q4NHT3	9MICC	Q4nht3	arthrobacte
599	29	85.3	231	2	Q82M12	STRAW	Q82m12	streptomyce	672	29	85.3	301	2	Q9KKY6	VIBCH	Q9kky6	vibrio chol
600	29	85.3	232	2	Q5YQZ4	NOCPA	Q5yqz4	nocardia fa	673	29	85.3	301	2	Q86XV4	HUMAN	Q86xv4	homo sapien
601	29	85.3	232	2	Q9HTS3	PSEAE	Q9hts3	pseudomonas	674	29	85.3	305	2	Q7W2V9	BORBR	Q7w2v9	bordetella
602	29	85.3	233	2	Q8POPO	XANAC	Q8ppo0	xanthomonas	675	29	85.3	307	2	O7W2V9	BORBR	O7w2v9	bordetella
603	29	85.3	236	2	Q57YF1	9TRYP	Q57yf1	trypanosoma	676	29	85.3	307	2	O7W2V9	BORBR	O7w2v9	bordetella
604	29	85.3	238	2	Q835J7	ENTFA	Q835j7	enterococcu	677	29	85.3	307	2	O7W2V9	BORBR	O7w2v9	bordetella
605	29	85.3	238	2	Q67QU0	SYMTH	Q67qu0	syndiobacte	678	29	85.3	308	2	O7W2V9	BORBR	O7w2v9	bordetella
606	29	85.3	239	2	Q92387	TRIVR	Q92387	trigonopsis	679	29	85.3	308	2	O7W2V9	BORBR	O7w2v9	bordetella
607	29	85.3	239	2	Q4LS99	9BURK	Q4ls99	burkholderi	680	29	85.3	308	2	O7W2V9	BORBR	O7w2v9	bordetella
608	29	85.3	239	2	Q7WID9	BORPA	Q7wid9	bordetella	681	29	85.3	309	2	O7W2V9	BORBR	O7w2v9	bordetella
609	29	85.3	241	2	Q7P147	CHRV0	Q7p147	chromobacte	682	29	85.3	309	2	O7W2V9	BORBR	O7w2v9	bordetella
610	29	85.3	242	2	Q8TXS6	METKA	Q8txs6	methanopyru	683	29	85.3	309	2	O7W2V9	BORBR	O7w2v9	bordetella
611	29	85.3	243	2	Q45899	COXBU	Q45899	coxiella bu	684	29	85.3	311	1	GPDA	HELHP	GPDA	helicobacte
612	29	85.3	243	2	Q45925	COXBU	Q45925	coxiella bu	685	29	85.3	311	2	O8XRE2	RALSO	O8xre2	raistonia s
613	29	85.3	243	2	Q9S618	COXBU	Q9s618	coxiella bu	686	29	85.3	313	2	Q4HVD0	GIBZE	Q4hvd0	gibberella
614	29	85.3	243	2	Q45943	COXBU	Q45943	coxiella bu	687	29	85.3	315	2	Q4GYV8	9TRYP	Q4gyv8	trypanosoma
615	29	85.3	244	2	Q7W0F8	BORPE	Q7w0f8	bordetella	688	29	85.3	318	2				

689	29	85.3	319	2	Q9RKM5_STRCO	Q9rkms streptomyce	762	1	MSRAB_VIBCH	99klx6 v peptide m	
690	29	85.3	324	2	Q5L6A8_CHLAB	Q5l6a8 chlamydophi	763	382	1	TGFBI_ONCMY	Q93449 oncorhynchu
691	29	85.3	324	2	Q5QX30_IDILO	Q5qx30 idiomarina	764	382	2	Q7V7J6_PROMM	Q7v7j6 prochloroco
692	29	85.3	326	1	YGGR_ECOLI	P52052 escherichia	765	383	2	Q827H5_STRAW	Q827h5 streptomyce
693	29	85.3	326	2	Q7NJV2_GLOVI	Q7njv2 gloeobacter	766	383	1	BODG_RHILO	Q98k0 rhizobium l
694	29	85.3	326	2	Q8JPA4_ROTTHU	Q8jpa4 human rotav	767	384	2	Q6PDD6_MOUSE	Q6pdd6 mus musculu
695	29	85.3	326	2	Q9M174_GREOV	Q9m174 human rotav	768	385	2	Q53CW4_9GAMA	Q53cw4 macaca fusc
696	29	85.3	326	2	Q9M176_GREOV	Q9m176 human rotav	769	386	2	Q86S16_HUMAN	Q86s16 homo sapien
697	29	85.3	327	2	Q8RS23_STRVL	Q8rs23 streptomyce	770	390	2	Q5Y289_NOCPA	Q5y289 nocardia fa
698	29	85.3	327	2	Q9KPN2_VIBCH	Q9kpn2 vibrio chol	771	391	2	Q6UHH2_ORYSA	Q6uuh2 oryza sativ
699	29	85.3	327	2	Q9ZRR1_RHIME	Q9zrr1 rhizobium m	772	392	2	Q54FP2_DICDI	Q54fp2 dictyosteli
700	29	85.3	328	2	Q8WHN3_9ASTR	Q8whn3 xeranthemum	773	392	2	Q6NY63_BRARE	Q6ny63 brachydanio
701	29	85.3	329	2	Q6ALK4_HELVI	Q6alk4 heliothis v	774	393	2	Q5L8Z1_BACFN	Q5l8z1 bacteroides
702	29	85.3	332	2	Q78116_NEUCR	Q78116 neurospora	775	393	2	Q64P54_BACFR	Q64p54 bacteroides
703	29	85.3	337	2	Q83PN9_SHIFL	Q83pn9 shigella fl	776	393	2	Q8A0W8_BACTN	Q8a0w8 bacteroides
704	29	85.3	337	2	Q8XDC2_ECO57	Q8xdc2 escherichia	777	395	2	Q5FQZ5_GLUOX	Q5fqz5 gluconobact
705	29	85.3	338	1	LACD_LISIN	Q92ab7 listeria in	778	396	1	PORA_PYREFU	Q51804 pyrococcus
706	29	85.3	338	1	LACD_LISMF	Q723b1 listeria in	779	397	2	Q5V4T9_HALMA	Q5v4t9 haloarcula
707	29	85.3	338	1	LACD_LISMO	Q8y919 listeria mo	780	399	2	Q9SSD0_ARATH	Q9s8d0 arabidopsis
708	29	85.3	338	2	Q8R5T1_THETN	Q8r5t1 thermoanaer	781	405	1	IPSP_MOUSE	P70458 mus musculu
709	29	85.3	339	1	RIBD_MYCTU	P71677 m riboflavi	782	405	2	Q8N3Z3_HUMAN	Q8n3z3 homo sapien
710	29	85.3	339	2	Q4YZR6_PLABE	Q4yzz6 plasmodium	783	405	2	Q8BKC4_MOUSE	Q8bkc4 mus musculu
711	29	85.3	339	2	Q7U043_MYCBO	Q7u043 m probable	784	405	2	Q66JR2_MOUSE	Q66jr2 mus musculu
712	29	85.3	339	2	Q9CCP8_MYCLE	Q9ccp8 mycobacteri	785	405	2	Q5BKQ8_MOUSE	Q5bkq8 mus musculu
713	29	85.3	341	2	Q741P5_MYCPA	Q741p5 mycobacteri	786	405	2	Q8BU50_MOUSE	Q8bu50 m mus muscu
714	29	85.3	341	2	Q8XCJ9_ECO57	Q8xcj9 escherichia	787	405	2	Q8BVN1_MOUSE	Q8bvn1 mus musculu
715	29	85.3	341	2	Q83J83_SHIFL	Q83j83 shigella fl	788	406	2	Q88292_RAT	Q88292 rattus norv
716	29	85.3	343	2	Q68E76_AERPU	Q68e76 aeromonas p	789	406	2	Q66HL5_RAT	Q66hl5 rattus norv
717	29	85.3	343	2	Q6FB21_ACIAD	Q6fb21 acinetobact	790	408	1	Y2112_PYRAE	Q8zvuv9 pyrobaculum
718	29	85.3	344	2	Q9HN36_HALSA	Q9hn36 halobacteri	791	408	2	Q92NV5_RHIME	Q92nv5 rhizobium m
719	29	85.3	344	2	Q55829_9FLAV	Q55829 sokuluk vir	792	409	2	Q6C525_YARLI	Q6c525 yarrowia li
720	29	85.3	345	2	Q52LAI_MOUSE	Q52lai mus musculu	793	410	2	Q4GYV2_9TRYP	Q4gyv2 trypanosoma
721	29	85.3	347	2	Q9GSY4_AGAIM	Q9gsy4 agama impal	794	411	2	Q7VFN5_HBLHP	Q7vfn5 helicobacte
722	29	85.3	350	1	MUTY_SALTY	Q95869 salmonella	795	413	1	DXR_NITEU	Q82u01 nitrosomona
723	29	85.3	350	2	Q57K05_SALCH	Q57k05 salmonella	796	414	1	Y4VG_RHISN	Q53215 rhizobium s
724	29	85.3	350	2	Q8Z3U0_SALTI	Q8z3u0 salmonella	797	414	2	Q8UDV4_AGR75	Q8udv4 agrobacteri
725	29	85.3	350	2	Q5PMW2_SALPA	Q5pmw2 salmonella	798	415	2	Q4HK12_CAMLA	Q4hk12 campylobact
726	29	85.3	350	2	Q4S273_TETNG	Q4s273 tetracodon n	799	416	1	Y997_ARCFU	Q29265 archaeoglob
727	29	85.3	351	2	Q5AZ00_EMENI	Q5az00 aspergillus	800	416	2	Q9EXQ4_CAMJE	Q9exq4 campylobact
728	29	85.3	351	2	Q4HYD8_GIBZE	Q4hyd8 gibberella	801	416	2	Q4HHY9_CAMCO	Q4hh9 campylobact
729	29	85.3	351	2	Q95G10_USEUD	Q95g10 saxifragell	802	416	2	Q5HVA5_CAMJR	Q5hva5 campylobact
730	29	85.3	351	2	Q8FE29_ECOL6	Q8fe29 escherichia	803	416	2	Q822P5_CHLCV	Q822p5 chlamydomphi
731	29	85.3	352	1	ENGCI_VIBVU	Q8dcv7 vibrio vuln	804	416	2	Q9PPN1_CAMJE	Q9ppn1 campylobact
732	29	85.3	352	1	ENGCI_VIBVY	Q7mgz6 vibrio vuln	805	416	2	Q6ANX8_DESPS	Q6anx8 desulfotale
733	29	85.3	352	1	MATK_SAXCE	Q33078 saxifraga c	806	417	1	TRMU_YEAST	Q12093 saccharomyc
734	29	85.3	353	2	Q66F15_YERPS	Q66f15 versinia ps	807	417	2	Q6IUF4_YEAST	Q6iuf4 saccharomyc
735	29	85.3	353	2	Q8ZJ35_YERPE	Q8zj35 versinia pe	808	418	2	Q4HRQ8_CAMUP	Q4hrq8 campylobact
736	29	85.3	355	2	Q75A37_ASHGO	Q75a37 ashbya goss	809	423	2	Q8X513_ECO57	Q8x513 escherichia
737	29	85.3	355	2	Q60DU2_ORYSA	Q60du2 oryza sativ	810	425	2	Q9NPN3_HUMAN	Q9npn3 homo sapien
738	29	85.3	356	2	Q4LTP6_9BURK	Q4ltp6 burkholderi	811	425	2	Q83X64_STRRO	Q83x64 streptomyce
739	29	85.3	356	2	Q67JN9_SYMTH	Q67jn9 symbiobacte	812	426	1	YJIN_ECOLI	P39385 escherichia
740	29	85.3	357	2	Q8GX19_ARATH	Q8gx19 arabidopsis	813	426	2	Q659Y0_ECOLI	Q659y0 escherichia
741	29	85.3	357	2	Q9HY94_PSEAE	Q9hy94 pseudomonas	814	427	2	Q8G3V7_BIFLO	Q8g3v7 bifidobacte
742	29	85.3	359	1	ID12_METJA	Q59272 methanococc	815	427	2	Q8KG31_CHLTE	Q8kg31 chlorobium
743	29	85.3	360	2	Q8GLQ8_AERSA	Q8glq8 aeromonas s	816	433	2	Q86WN8_HUMAN	Q86wn8 homo sapien
744	29	85.3	361	2	Q62FF1_BURMA	Q62ff1 burkholderi	817	433	2	Q4H3B0_CTOIN	Q4h3b0 ciona intes
745	29	85.3	362	2	Q57U20_9TRYP	Q57u20 trypanosoma	818	434	2	Q7MZJ8_PHOLL	Q7mzj8 photorhabdu
746	29	85.3	362	2	Q60B22_METCA	Q60b22 methylococc	819	437	2	Q6JMF5_9NUCL	Q6jmf5 neodiprion
747	29	85.3	363	2	Q8TVS1_METKA	Q8tvsl methanopyru	820	438	2	Q8FA91_ECOL6	Q8fa91 escherichia
748	29	85.3	364	2	Q29447_ARCFU	Q29447 archaeoglob	821	439	2	Q6LSA2_ORYSA	Q6lsa2 oryza sativ
749	29	85.3	364	2	Q67SV1_SYMTH	Q67sv1 symbiobacte	822	440	2	Q30538_PSEAE	Q30538 pseudomonas
750	29	85.3	365	1	ID12_METKA	Q8tx99 methanopyru	823	440	2	Q8VWB7_9ACTO	Q8vwb7 streptomyce
751	29	85.3	367	2	Q5KF19_CRYNE	Q5kf19 cryptococcu	824	440	2	Q9LS55_9ACTO	Q9ls55 streptomyce
752	29	85.3	369	2	Q9U222_PYRAB	Q9u222 pyrococcus	825	440	2	Q9I330_PSEAE	Q9i330 pseudomonas
753	29	85.3	369	2	Q5RL65_MOUSE	Q5rl65 mus musculu	826	441	1	Y1340_HAEIN	P44165 haemophilus
754	29	85.3	370	2	Q8UJB9_AGR75	Q8ujb9 agrobacteri	827	445	2	Q70WP6_AERSA	Q70wp6 aeromonas s
755	29	85.3	371	2	Q4J8E9_SULAC	Q4j8e9 sulfolobus	828	449	1	Y942_SYNEL	Q8dkb7 synechococ
756	29	85.3	371	2	Q70SK6_9MYRT	Q70sk6 epilobium b	829	450	2	Q78CC9_NEUCR	Q78cc9 neurospora
757	29	85.3	371	2	Q70SK8_9MYRT	Q70sk8 epilobium c	830	450	2	Q500J0_PSESY	Q500j0 pseudomonas
758	29	85.3	371	2	Q98HB1_RHILO	Q98hb1 rhizobium l	831	451	1	Y1462_HAEIN	P45217 haemophilus
759	29	85.3	375	2	Q4NT18_9DELT	Q4nt18 anaeromoxob	832	454	2	Q4QKI6_HAE18	Q4qki6 haemophilus
760	29	85.3	376	2	Q86SF7_HUMAN	Q86sf7 homo sapien	833	455	2	Q603F8_METCA	Q603f8 methylococc
761	29	85.3	376	2	Q73TK1_MYCPA	Q73tk1 mycobacteri	834	455	2	Q8RH36_FUSNN	Q8rh36 fusobacteri

835	29	85.3	457	1	Y4BF_RHISN	P55373 rhizobium s	908	29	85.3	557	2	Q41IA7_GIBZE	Q41IA7 gibberella
836	29	85.3	457	2	Q8XV7_RALSO	Q8XV7 ralstonia s	909	29	85.3	558	2	Q69YK3_HUMAN	Q69YK3 homo sapien
837	29	85.3	458	2	Q55QW6_CRYNE	Q55QW6 cryptococcus	910	29	85.3	558	2	Q6FZX6_BARQU	Q6FZX6 bartonella
838	29	85.3	459	2	Q86W7_HUMAN	Q86W7 homo sapien	911	29	85.3	558	2	Q6G3A3_BARHE	Q6G3A3 bartonella
839	29	85.3	460	2	Q6MPFO_DBEBA	Q6MPFO dbellovibri	912	29	85.3	558	2	Q9CX72_MOUSE	Q9CX72 mus musculu
840	29	85.3	464	2	Q4J5N0_AZOBAC	Q4J5N0 azotobacter	913	29	85.3	558	2	Q8BM95_MOUSE	Q8BM95 mus musculu
841	29	85.3	464	2	Q4J5N0_AZOV1	Q4J5N0 streptomyce	914	29	85.3	560	2	Q51X3_MAGGR	Q51X3 magnaporthe
842	29	85.3	464	2	Q54204_STRCO	Q54204 streptomyce	915	29	85.3	562	2	Q8DMI6_SYNEL	Q8DMI6 synechococc
843	29	85.3	465	1	MPBP_BLAEM	Q03032 blastocladi	916	29	85.3	564	2	Q8QS39_9BETA	Q8QS39 pongine her
844	29	85.3	466	2	Q5NZR1_AZOSE	Q5NZR1 azoarcus sp	917	29	85.3	567	1	G6PIA_ORYSA	G6PIA oryza sativ
845	29	85.3	466	2	Q82JF1_STRAW	Q82JF1 streptomyce	918	29	85.3	567	1	G6PIA_MAIZE	G6PIA maize
846	29	85.3	467	2	Q4Q160_LEIMA	Q4Q160 leishmania	919	29	85.3	567	2	Q86WN3_HUMAN	Q86WN3 homo sapien
847	29	85.3	467	2	Q5E6K2_VIBF1	Q5E6K2 vibrio fisc	920	29	85.3	567	2	Q5CTF8_CRYPV	Q5CTF8 cryptospori
848	29	85.3	470	2	Q7TM95_MOUSE	Q7TM95 mus musculu	921	29	85.3	567	2	Q5CKR9_CRYHO	Q5CKR9 cryptospori
849	29	85.3	471	2	Q7W4T3_BORDE	Q7W4T3 bordetella	922	29	85.3	567	2	Q8H3D6_ORYSA	Q8H3D6 oryza sativ
850	29	85.3	472	2	Q8TZL5_PYRFU	Q8TZL5 pyrococcus	923	29	85.3	567	2	Q652G2_ORYSA	Q652G2 oryza sativ
851	29	85.3	473	2	Q7PKB4_ANOGA	Q7PKB4 anopheles g	924	29	85.3	567	2	Q8H8M6_ORYSA	Q8H8M6 oryza sativ
852	29	85.3	473	2	Q8KL89_RHIET	Q8KL89 rhizobium e	925	29	85.3	567	2	Q41PB0_GIBZE	Q41PB0 gibberella
853	29	85.3	474	2	Q7VU97_BORPE	Q7VU97 bordetella	926	29	85.3	568	2	Q94Z24_PYLII	Q94Z24 pylaiaella 1
854	29	85.3	474	2	Q7WGA9_BORBR	Q7WGA9 bordetella	927	29	85.3	568	2	Q6AH33_LEIXX	Q6AH33 leifsonia x
855	29	85.3	474	2	Q72AD1_DSUVH	Q72AD1 desulfovibr	928	29	85.3	570	1	UL49_HCMVA	UL49 hcmva
856	29	85.3	475	2	Q9KWD8_AGRRH	Q9KWD8 agrobacteri	929	29	85.3	570	1	Q6RXG9_HCMV	Q6RXG9 human cytom
857	29	85.3	479	2	Q55IU1_CRYNE	Q55IU1 cryptococcus	930	29	85.3	570	2	Q6SW82_HCMV	Q6SW82 human cytom
858	29	85.3	479	2	Q5KCM0_CRYNE	Q5KCM0 cryptococcus	931	29	85.3	572	2	Q80FH8_9LUTE	Q80FH8 sugarcane y
859	29	85.3	482	2	Q9JPA7_RHOGE	Q9JPA7 rhodocyclu	932	29	85.3	572	2	Q91L47_9LUTE	Q91L47 sugarcane y
860	29	85.3	482	2	Q5WND6_CAEBR	Q5WND6 caenorhabd	933	29	85.3	572	2	Q91L51_9LUTE	Q91L51 sugarcane y
861	29	85.3	484	2	Q83YGB_RHOGE	Q83YGB rhodocyclu	934	29	85.3	572	2	Q91L59_9LUTE	Q91L59 sugarcane y
862	29	85.3	485	2	Q5QD39_HUMAN	Q5QD39 homo sapien	935	29	85.3	572	2	Q91H78_9LUTE	Q91H78 sugarcane y
863	29	85.3	485	2	Q916T6_PBEAE	Q916T6 pseudomonas	936	29	85.3	572	2	Q9QON8_9LUTE	Q9QON8 sugarcane y
864	29	85.3	491	2	Q86WN6_HUMAN	Q86WN6 homo sapien	937	29	85.3	572	2	Q72PW8_LEPIC	Q72PW8 leptospira
865	29	85.3	492	2	Q86SJ3_HUMAN	Q86SJ3 homo sapien	938	29	85.3	574	2	Q8F6B9_LEPIN	Q8F6B9 leptospira
866	29	85.3	497	2	Q4P263_USTMA	Q4P263 ustilago ma	939	29	85.3	575	2	Q4PFD2_USTMA	Q4PFD2 ustilago ma
867	29	85.3	501	2	Q8H336_HUMAN	Q8H336 homo sapien	940	29	85.3	576	2	Q7U218_MYCBO	Q7U218 mycobacteri
868	29	85.3	501	2	Q6N6Q2_RHOPA	Q6N6Q2 rhodopseudo	941	29	85.3	576	2	P96408_MYCTU	P96408 mycobacteri
869	29	85.3	502	2	Q98862_9ASTE	Q98862 petalonyx n	942	29	85.3	576	2	Q96RH2_BRAJA	Q96RH2 bradyrhizob
870	29	85.3	504	2	Q4UF69_THEAN	Q4UF69 theileria a	943	29	85.3	580	2	1_PDE9A_HUMAN	1_PDE9A homo sapien
871	29	85.3	504	2	Q8GW59_ARATH	Q8GW59 arabidopsis	944	29	85.3	593	1	PDE9A_HUMAN	PDE9A homo sapien
872	29	85.3	504	2	Q6X5V5_9ASTE	Q6X5V5 petalonyx l	945	29	85.3	595	2	Q7SCV2_NEUCR	Q7SCV2 neurospora
873	29	85.3	505	2	Q8N3F2_HUMAN	Q8N3F2 homo sapien	946	29	85.3	595	2	Q4WAU8_ASPFU	Q4WAU8 aspergillus
874	29	85.3	505	2	Q7UG18_RHOBA	Q7UG18 rhodopirell	947	29	85.3	597	2	Q64753_ARATH	Q64753 arabidopsis
875	29	85.3	506	2	Q8MD40_9ASTE	Q8MD40 petalonyx l	948	29	85.3	597	2	Q64753_ARATH	Q64753 arabidopsis
876	29	85.3	506	2	Q8MD41_9ASTE	Q8MD41 petalonyx t	949	29	85.3	598	2	Q817N1_CAEEL	Q817N1 caenorhabd
877	29	85.3	507	2	Q8MD43_9ASTE	Q8MD43 petalonyx p	950	29	85.3	601	2	Q4SU34_TETNG	Q4SU34 tetraodon n
878	29	85.3	507	2	Q4WL22_ASPFU	Q4WL22 aspergillus	951	29	85.3	603	2	Q5E2P0_VIBF1	Q5E2P0 vibrio fisc
879	29	85.3	507	2	Q86WN9_HUMAN	Q86WN9 homo sapien	952	29	85.3	607	1	GLMS_CLOTE	GLMS clostridia
880	29	85.3	510	2	Q6C469_YARLI	Q6C469 yarrowia li	953	29	85.3	607	2	Q51A19_ENTHI	Q51A19 entamoeba h
881	29	85.3	510	2	Q6BPB4_DEBHA	Q6BPB4 debaryomyce	954	29	85.3	607	2	Q4H636_9DEIO	Q4H636 deinococcus
882	29	85.3	517	2	Q6BPB4_DEBHA	Q6BPB4 debaryomyce	955	29	85.3	608	2	Q5U483_XENLA	Q5U483 xenopus lae
883	29	85.3	526	2	Q86WN5_HUMAN	Q86WN5 homo sapien	956	29	85.3	611	2	Q5YRV8_NOCFA	Q5YRV8 nocardia fa
884	29	85.3	530	2	Q5P026_AZOSE	Q5P026 azoarcus sp	957	29	85.3	613	2	Q4IQ02_GIBZE	Q4IQ02 gibberella
885	29	85.3	531	2	Q87WE7_PSESM	Q87WE7 pseudomonas	958	29	85.3	615	2	Q4S3B0_TETNG	Q4S3B0 tetraodon n
886	29	85.3	533	2	Q53Y40_HUMAN	Q53Y40 homo sapien	959	29	85.3	615	2	Q8K297_MOUSE	Q8K297 mus musculu
887	29	85.3	534	1	PDE9A_MOUSE	Q70628 mus musculu	960	29	85.3	617	2	Q6PGL1_MOUSE	Q6PGL1 mus musculu
888	29	85.3	534	2	Q8BSU4_MOUSE	Q8BSU4 rattus norv	961	29	85.3	619	1	SC6A3_MOUSE	SC6A3 mus musculu
889	29	85.3	534	2	Q8BSU4_MOUSE	Q8BSU4 rattus norv	962	29	85.3	619	1	SC6A3_MOUSE	SC6A3 mus musculu
890	29	85.3	535	2	Q6EUF2_ORYSA	Q6EUF2 oryza sativ	963	29	85.3	619	1	SC6A3_MOUSE	SC6A3 mus musculu
891	29	85.3	537	2	Q871H6_NEUCR	Q871H6 neurospora	964	29	85.3	619	1	SC6A3_MOUSE	SC6A3 mus musculu
892	29	85.3	537	2	Q13369_ASCIM	Q13369 ascobolus i	965	29	85.3	619	2	Q8C009_MOUSE	Q8C009 mus musculu
893	29	85.3	540	1	BB61_RABIT	Q05004 oryctolagus	966	29	85.3	619	2	Q8C009_MOUSE	Q8C009 mus musculu
894	29	85.3	540	2	Q86WP0_HUMAN	Q86WP0 homo sapien	967	29	85.3	620	1	SC6A3_HUMAN	SC6A3 mus musculu
895	29	85.3	541	2	Q8YL89_ANASP	Q8YL89 anabaena sp	968	29	85.3	620	1	SC6A3_HUMAN	SC6A3 mus musculu
896	29	85.3	542	2	Q5BDP1_EMENI	Q5BDP1 aspergillus	969	29	85.3	620	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
897	29	85.3	542	2	Q4WCK6_ASPFU	Q4WCK6 aspergillus	970	29	85.3	620	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
898	29	85.3	543	2	Q4PB4_USTMA	Q4PB4 ustilago ma	971	29	85.3	620	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
899	29	85.3	543	2	Q4IBN4_GIBZE	Q4IBN4 gibberella	972	29	85.3	620	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
900	29	85.3	543	2	Q5GV08_XANOR	Q5GV08 xanthomonas	973	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
901	29	85.3	543	2	Q4V062_XANCP	Q4V062 xanthomonas	974	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
902	29	85.3	544	2	Q8PDX0_XANCP	Q8PDX0 xanthomonas	975	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
903	29	85.3	544	2	Q7S026_NEUCR	Q7S026 neurospora	976	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
904	29	85.3	549	2	Q7SFQ9_NEUCR	Q7SFQ9 neurospora	977	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
905	29	85.3	549	2	Q529E4_ORYSA	Q529E4 oryza sativ	978	29	85.3	622	2	Q7JHP8_MACFA	Q7JHP8 macaca fasc
906	29	85.3	551	2	Q92PP7_RHIME	Q92PP7 rhizobium m	979	29	85.3	625	2	Q9GJT5_SALSC	Q9GJT5 salmirel sci
907	29	85.3	552	2	Q9KX16_STRCO	Q9KX16 streptomyce	980	29	85.3	629	2	Q9GJT5_SALSC	Q9GJT5 salmirel sci
					Q86WN4_homo sapien								

981 29 85.3 633 1 NUOL MYCTU
982 29 85.3 633 2 Q7TX55 MYCBO
983 29 85.3 639 2 Q8TH4 MYTAC
984 29 85.3 641 2 Q4QCK1 LEIMA
985 29 85.3 641 2 Q54MD1 D1CDDI
986 29 85.3 645 2 Q8C011 MOUSE
987 29 85.3 649 2 Q5ST66 CRYNE
988 29 85.3 649 2 Q5KI22 CRYNE
989 29 85.3 651 2 Q8T749 BRAFL
990 29 85.3 653 1 Y1364 MYCTU
991 29 85.3 653 2 Q7U068 MYCBO
992 29 85.3 654 2 Q5FQ22 GLUOX
993 29 85.3 657 2 Q61X34 CAEBR
994 29 85.3 662 2 Q5PAC8 ANAMM
995 29 85.3 663 2 Q7QF31 ANOXA
996 29 85.3 674 2 Q5ST30 CRYNE
997 29 85.3 674 2 Q5KHG3 CRYNE
998 29 85.3 677 2 Q8BNT0 MOUSE
999 29 85.3 678 2 Q7UWE7 RHOB
1000 29 85.3 682 2 Q4SHT0 TETNG

ALIGNMENTS

RESULT 1
ID Q8Z285 PYRAE PRELIMINARY; PRT; 78 AA.
AC Q8Z285;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein PAE0388.
GN OrderedLocusNames=PAE0388;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AE009765; AAL62756.1; -, Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 78 AA; 8694 MW; 382A6C0C24D130D3 CRC64;
Query Match 100.0%; Score 34; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 17 TLREWL 22
RESULT 2
ID Q5OWB5 ENTHI PRELIMINARY; PRT; 78 AA.
AC Q5OWB5;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=173.t00016;
OS Entamoeba histolytica HM-1.1MS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1.1MS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moulé S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gichrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichteritz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica";
RL Nature 433:865-868 (2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAPB01000543; EAL45891.1; -, Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 78 AA; 8632 MW; 3EA7ED258C3690F4 CRC64;
Query Match 100.0%; Score 34; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 52 TLREWL 57
RESULT 3
ID Q9SHP3 ARATH PRELIMINARY; PRT; 190 AA.
AC Q9SHP3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE FIK23.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Chao Q., Shinn P., Brooks S., Buehler E., Dunn P., Khan S., Kim C.,
RA Walker M., Brooks S., Altafi H., Araujo R., Conn L., Conway A.B.,
RA Gonzalez A., Hansen N.P., Huizar L., Kremenetskaia I., Lenz C., Li J.,
RA Liu S., Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V.,
RA Yu G., Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC Cheuk R., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.


```

[4]
RP NUCLEOTIDE SEQUENCE.
RA Shin P., Brooks S., Buehler E., Chao Q., Cheuk R., Johnson-Hopson C.,
RA Khan S., Kim C., Alkafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.; to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007508; AAF24553.1; -; Genomic DNA.
SQ SEQUENCE 190 AA; 22057 MW; EA7455FDC49B0358 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
   |||||
Db 15 TLREWL 20

RESULT 4
Q4T3D3 TETNG PRELIMINARY; PRT; 213 AA.
AC Q4T3D3;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome undetermined SCAF10097, whole genome shotgun sequence.
GN ORFNames=GSTENG0007901001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Foulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB010097; CAF92599.1; -; Genomic DNA.
SQ SEQUENCE 213 AA; 23088 MW; 23BBE3F0FF069D82 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
   |||||
Db 53 TLREWL 58

```

```

RESULT 5
Q7TNX8 RAT PRELIMINARY; PRT; 243 AA.
AC Q7TNX8;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE LRP16-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SD;
RA Guo J.H., Yu L.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF419856; AAP97291.1; -; mRNA.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR SMART; SM00506; Alpp; 1.
SQ SEQUENCE 243 AA; 27290 MW; D0F593336EDD3BEF CRC64;

Query Match 100.0%; Score 34; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
   |||||
Db 204 TLREWL 209

RESULT 6
LRP16 RAT STANDARD; PRT; 258 AA.
AC LRK4G6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Protein LRP16 (Fragment).
GN Name=Lrp16;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley;
RA Zhang X., Ip N.Y.;
RT "Rat mRNA sequence similar to LRP16 protein of humans.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- MISCELLANEOUS: Overexpression may promote MCF-7 cells
CC proliferation (By similarity).
CC -!- SIMILARITY: Contains 1 Alpp domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AF404762; AAM45760.1; ALT_INIT; mRNA.
DR Ensembl; ENSRNOG00000021174; Rattus norvegicus.
DR RGD; 628701; Lrp16.
DR InterPro; IPR002589; Alpp.
DR Pfam; PF01661; Alpp; 1.
DR SMART; SM00506; Alpp; 1.
FT DOMAIN 103..217 Alpp.
FT NON_TER 1
SQ SEQUENCE 258 AA; 28643 MW; FB430516A12D6B42 CRC64;

```

```
Query Match      100.0%; Score 34; DB 1; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 219 TLREWL 224

RESULT 7
KSGA AZOSE STANDARD; PRT; 265 AA.
ID KSGA AZOSE STANDARD; PRT; 265 AA.
AC QSP7JJ;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Dimethyladenosine transferase (EC 2.1.1.-) (S'-adenosylmethionine-6-N',
DE N'-adenosyl(rRNA) dimethyltransferase) (16S rRNA dimethylase) (High
DE level kasugamycin resistance protein ksgA) (Kasugamycin
DE dimethyltransferase).
GN Name=ksgA; OrderedLocusNames=AZOSEA05980; ORFNames=ebA1138;
OS Azotarcus sp. (strain EBN1).
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azotarcus.
OX NCBI_TaxID=76114;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=15551059; DOI=10.1007/s00203-004-0742-9;
RX Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel P.,
RA Reinhardt R.
RT "The genome sequence of an anaerobic aromatic-degrading denitrifying
RT bacterium, strain EBN1."
RL Arch. Microbiol. 183:27-36(2005).
CC -1- FUNCTION: Specifically dimethylates two adjacent adenosines in the
CC loop of a conserved hairpin near the 3' end of 16S rRNA in the 30S
CC particle. Its inactivation leads to kasugamycin resistance (by
CC similarity).
CC -1- SIMILARITY: Belongs to the rRNA adenine N-6-methyltransferase
CC family. ksgA subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; CR555306; CAI06720.1; -; Genomic_DNA.
CC HAMAP; MF 00607; -; 1.
CC InterPro; IPR011530; ksgA.
CC InterPro; IPR001737; RNA_meth_trans.
CC InterPro; IPR000051; SAM_Bd.
CC PANTHER; PTHR11727:SF6; ksgA; 1.
CC PANTHER; PTHR11727; RNA_meth_trans; 1.
CC Pfam; PF00398; RnaAD; 1.
CC SMART; SM00650; YADC; 1.
CC TIGRFAMs; TIGR00755; ksgA; 1.
CC PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Complete proteome; Methyltransferase;
KW rRNA processing; Transferase.
SQ SEQUENCE 265 AA; 29742 MW; 239359142EAD5597 CRC64;

Query Match      100.0%; Score 34; DB 1; Length 265;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 218 TLREWL 223

RESULT 8
Q87IG5 VIBPA PRELIMINARY; PRT; 277 AA.
ID Q87IG5 VIBPA PRELIMINARY; PRT; 277 AA.
AC Q87IG5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transcriptional regulator, LysR family.
GN OrderedLocusNames=VPA0641;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=RIMD 2210633 / Serotype O3:k6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
RX EMBL; BA000032; BAC61984.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN_1.
DR PROSITE; PS00931; HTH_LYSR; 1.
KW Complete proteome.
SQ SEQUENCE 277 AA; 30745 MW; A7E211E9E24F23C2 CRC64;

Query Match      100.0%; Score 34; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 129 TLREWL 134

RESULT 9
Q9KX21 OLICA PRELIMINARY; PRT; 296 AA.
ID Q9KX21 OLICA PRELIMINARY; PRT; 296 AA.
AC Q9KX21;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CoxK.
GN Name=CoxK;
OS Oligotropha carboxidovorans (Pseudomonas carboxydovorans).
OG Plasmid pHC3.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Oligotropha.
OX NCBI_TaxID=40137;
RN [1]
NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RX MEDLINE=95238294; PubMed=7721710;
RA Schuebel U., Kraut M., Moersdorf G., Meyer O.;
RT "Molecular characterization of the gene cluster encoding the
RT molybdenum-containing carbon monoxide dehydrogenase of Oligotropha
RT carboxidovorans."
RL J. Bacteriol. 177:2197-2203(1995).
RN [2]
NUCLEOTIDE SEQUENCE.
RP STRAIN=OM5;
RX MEDLINE=97464431; PubMed=9324252;
RA Santiago B., Meyer O.;
```

RT "Purification and molecular characterization of the H2 uptake
 RT membrane-bound Nife-hydrogenase from the carboxidotrophic bacterium
 RT Oligotropha carboxidovorans."; J. Bacteriol. 179:6053-6060(1997).
 RL [3]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=OM5;
 RC Santiago B., Schuebel U., Egelseer C., Meyer O.;
 RA "Sequence analysis, characterization and CO-specific transcription of
 RT the cox gene cluster on the megaplasmid pHCG3 of Oligotropha
 RT carboxidovorans."; Gene 236:1157-1247(1999).
 RL [4]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=OM5;
 RC PubMed=1464498; DOI=10.1016/j.gene.2003.08.027;
 RX Fuhrmann S., Ferner M., Jeffke T., Henne A., Gottschalk G., Meyer O.;
 RA "Complete nucleotide sequence of the circular megaplasmid pHCG3 of
 RT Oligotropha carboxidovorans: function in the chemolithoautotrophic
 RT utilization of CO, H(2) and CO(2)."; Gene 322:67-75(2003).
 RL [5]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=OM5;
 RC Schuebel U.;
 RA Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RL [6]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=OM5;
 RC Santiago B.;
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL [7]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=OM5;
 RC Fuhrmann S.;
 RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; X82447; CAB76251.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR000620; DUF6_TM.
 DR Pfam; PF00892; DUF6; 2.
 KW Plasmid.
 SQ SEQUENCE 296 AA; 30979 MW; 174A75A5796CE9F2 CRC64;
 Query Match 100.0%; Score 34; DB 2; Length 296;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 74 TLREWL 79
 RESULT 10
 LRP16 HUMAN
 ID LRP16_HUMAN STANDARD; PRT; 325 AA.
 AC Q9BQ69; Q9UH96;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Protein LRP16.
 GN Name=LRP16;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Lymphocyte;
 RX PubMed=12578638;
 RA Han W.-D., Yu L., Lou F.D., Wang Q.S., Zhao Y., Shi Z.J., Jin H.J.;
 RT "The application of RACE technique to clone the full-length cDNA of a
 RT novel leukemia associated gene LRP16.";

RL Zhongguo Shi Yan Xue Ye Xue Za Zhi 9:18-21(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Brain, and Eye;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fhey J.J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
 RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Jones S.J.M., Smailus D.E.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Maria M.A.;
 RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [3]
 RN INDUCTION.
 RP MEDLINE=22675902; PubMed=12790785; DOI=10.1677/erc.0.0100217;
 RA Han W.-D., Mu Y.-M., Lu X.-C., Xu Z.-M., Li X.-J., Yu L., Song H.-J.,
 RA Li M., Lu J.-M., Zhao Y.-L., Pan C.-Y.;
 RT "Up-regulation of LRP16 mRNA by 17beta-estradiol through activation of
 RT estrogen receptor alpha (ERalpha), but not ERbeta, and promotion of
 RT human breast cancer MCF-7 cell proliferation: a preliminary report.";
 RL Endocr. Relat. Cancer 10:217-224(2003).
 CC -I- INDUCTION: Overexpressed by estrogens in breast cancer MCF-7
 CC cells, probably via an activation of nuclear receptors for
 CC steroids (ESR1 and ESR2).
 CC -I- MISCELLANEOUS: Overexpression may promote MCF-7 cells
 CC proliferation.
 CC -I- SIMILARITY: Contains 1 Alpp domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF202922; AAF15294.2; -; mRNA.
 DR EMBL; BC000270; AAH00270.2; -; mRNA.
 DR EMBL; BC003188; AAH03188.1; ALT INIT; mRNA.
 DR EMBL; BC007297; AAH07297.1; -; mRNA.
 DR EMBL; BC008316; AAH08316.1; -; mRNA.
 DR HSSP; O28751; IHU2.
 DR Ensembl; ENSG00000133315; Homo sapiens.
 DR H-InvDB; HIX0019413; -.
 DR InterPro; IPR002589; Alpp.
 DR Pfam; PF01661; Alpp; 1.
 FT DOMAIN 170 284.
 FT SEQUENCE 325 AA; 35505 MW; 82294BFC904FA4D0 CRC64;
 SQ
 Query Match 100.0%; Score 34; DB 1; Length 325;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 286 TLREWL 291
 RESULT 11
 Q75PRO_PHYPA
 ID Q75PRO_PHYPA PRELIMINARY; PRT; 391 AA.
 AC Q75PRO;

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Vascular plant one zinc finger protein.
GN Name=PpVOZ1;
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariaceae; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mitsuoka N., Hisebori T., Takeyasu K., Sato M.H.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB164647; BAD12233.1; -; mRNA.
SQ SEQUENCE 391 AA; 42863 MW; 7849223F8A834861 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 210 TLREWL 215

RESULT 12
ID Q7PSJ0 ANOGA PRELIMINARY; PRT; 473 AA.
AC Q7PSJ0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000012388 (Fragment).
GN ORFNames=ENSANGG0000009899;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Anopheles gambiae Sequence Committee;
RC STRAIN=PEST;
RG "Anopheles gambiae re-annotation.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAB01008823; EAA05508.3; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 2.
DR SMART; SM00220; S_TKc_1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1 473
FT NON_TER 473 473

```

```

SQ SEQUENCE 473 AA; 53729 MW; A994231087902F7C CRC64;

Query Match 100.0%; Score 34; DB 2; Length 473;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 280 TLREWL 285

RESULT 13
ID Q5SGQ0 ARATH PRELIMINARY; PRT; 486 AA.
AC Q5SGQ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE P3M18.4 (Atg1g28520) (Transcription factor AtVOZ1).
GN Name=AtVOZ1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federepiel N.A., Theologis A.,
RA Ecker J.R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federepiel N., Theologis A.,
RA Ecker J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Mitsuoka N., Hisebori T., Takeyasu K., Sato M.H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010155; AAP16771.1; -; Genomic_DNA.
DR EMBL; BT020261; AAV84482.1; -; mRNA.
DR EMBL; AB252556; BAD17857.1; -; mRNA.
DR PIR; B86411; B86411.
SQ SEQUENCE 486 AA; 54079 MW; 9BA92951E2128858 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
Db 311 TLREWL 316

RESULT 14
ID Q5BGN6 EMENI PRELIMINARY; PRT; 522 AA.
AC Q5BGN6;

```

10-MAY-2005 (TrEMBLrel. 30, Created)
 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AN0294.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eukaryotes; Trichocomaceae; Emericella.
 OX NCBI_TaxID=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
 RA Boukhalil B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gherre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stangor-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Teafaye S., Theodore J., Topham K., Travers M.,
 RA Vasiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Zander E.;
 RT "Genome Sequence of Aspergillus nidulans.";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAC01000006; EAA65700.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 522 AA; 58714 MW; 58DE8D4C9696F385 CRC64;
 Query Match 100.0%; Score 34; DB 2; Length 522;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 200 TLREWL 205
 RESULT 15
 Q5XPC4 PHYN PRELIMINARY; PRT; 556 AA.
 AC Q5XPC4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Glucose-6-phosphate isomerase.
 GN Name=GPI;
 OS Phytophthora infestans (Potato late blight fungus).
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4787;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22961941; PubMed=14599887; DOI=10.1016/S1087-1845(03)00107-5;
 RA Ospina-Giraldo M.D., Jones R.W.;
 RT "Characterization of the glucose-6-phosphate isomerase gene in
 RT Phytophthora infestans reveals the presence of multiple alleles.";
 RL Fungal Genet. Biol. 40:197-206(2003)
 CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-

phosphate.
 CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the GPI family.
 DR EMBL; AV220242; AAP51062.1; -; Genomic DNA.
 DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; P:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isoomerase.
 DR Pfam; PF00342; GPI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P-GLUCOSE-ISOMERASE_1; 1.
 DR PROSITE; PS00174; P-GLUCOSE-ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 556 AA; 60840 MW; AACAC63EB68B752 CRC64;
 Query Match 100.0%; Score 34; DB 2; Length 556;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 221 TLREWL 226
 RESULT 16
 Q4HCY6_9DE10 PRELIMINARY; PRT; 557 AA.
 ID Q4HCY6_9DE10 PRELIMINARY; PRT; 557 AA.
 AC Q4HCY6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Hypothetical protein precursor.
 DE ORFNames=DgeodRAFT_1919;
 GN Deinococcus geothermalis DSM 11300.
 OS Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=319795;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DSM 11300;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
 RT DSM 11300.";
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DSM 11300;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Deinococcus geothermalis
 RT DSM 11300.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAHE01000001; EAL93813.1; -; Genomic_DNA.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 42 Potential.
 SQ SEQUENCE 557 AA; 60854 MW; 1FC20B945E00DF6F CRC64;
 Query Match 100.0%; Score 34; DB 2; Length 557;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 109 TLREWL 114

```

RESULT 17
ID Q5W7M1_XENTR PRELIMINARY; PRT; 597 AA.
AC Q5W7M1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Hypothetical LOC496849.
GN Name=LOC496849;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Anophelinae; Xenopus; Silurana.
OX NCBI_TaxID=81364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Fellngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiy S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC088562; AA088562.1; -; mRNA.
DR SRR; Q5W7M1; 302-530.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR SMART; SM00028; TPR; 6.
DR SMART; PS00005; TPR; 5.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 597 AA; 67162 MW; BF2D14AEAL1CB87 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 4.9e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 1 TLREWL 6
Db 410 TLREWL 415

RESULT 18
ID Q6FLE9_CANGA PRELIMINARY; PRT; 637 AA.
AC Q6FLE9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Similar to sp|P53925 Saccharomyces cerevisiae YNL15C.
GN OrderedLocusNames=CAGL0L03938g;
OS Candida glabrata (Yeast) (Torulopsis glabrata).

```

```

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons A., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenn S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
DR EMBL; CR380958; CAG61915.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000073; A/b_hydrolase.
DR Pfam; PF00561; Abhydrolase_1; 1.
KW Complete proteome.
SQ SEQUENCE 637 AA; 73944 MW; D107B616C63CB755 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 5.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 1 TLREWL 6
Db 261 TLREWL 266

RESULT 19
ID Q9VYJ0_DROME PRELIMINARY; PRT; 963 AA.
AC Q9VYJ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG32648-PA.
GN Name=Pde9; ORFNames=CG32648;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT melanogaster eukaryotic genome sequence;"
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celinker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective;"
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.K., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review;"
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley *Drosophila* Genome Project;
RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence;"
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RA Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -/- INTERACTION:
CC Q9VLM2:CG13384; NDEXP=1; IntAct=EBI-155878, EBI-109582;
DR EMBL; AE003490; AAF48205.2; -; Genomic_DNA.
DR HSSP; Q08499; IOYN.
DR IntAct; Q9VJ00; -.
DR Ensembl; CG32648; *Drosophila melanogaster*.

DR FlyBase; FBgn0052648; CG32648.
DR FlyBase; FBgn0052648; Pde9.
DR GO; GO:0004114; F:3'; 5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003607; Met_phos_hydro.
DR InterPro; IPR002073; PDease.
DR InterPro; IPR000217; Tubulin.
DR Pfam; PF00233; PDease_I; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR SMART; SM00471; HDC; 1.
DR SMART; SM00126; PDEASE_I; 1.
DR PROSITE; PS00227; TUBULIN; UNKNOWN 1.
DR PROSITE; PS00227; TUBULIN; UNKNOWN 1.
SQ SEQUENCE 963 AA; 103516 MW; B1987E0DF6B14F CRC64;
Query Match 100.0%; Score 34; DB 2; Length 963;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TLREWL 6
DB 93 TLREWL 98
RESULT 20
Q9UVCL CLAFU PRELIMINARY; PRT; 1243 AA.
ID Q9UVCL CLAFU PRELIMINARY;
AC Q9UVCL;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Cladosporium fulvum (Fulvia fulva).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothymycetes incertae sedis;
OC Mycosphaerellaceae; Mitosporic Mycosphaerellaceae; Cladosporium.
OX NCBI_TaxID=5499;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Simpson M.L., Butler M., Poulter R.T.M.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -/- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF051915; AAF21678.1; -; Genomic_DNA.
DR HSSP; P05205; 1KNA.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003682; F:chromatin binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006333; P:chromatin assembly or disassembly; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; RVTse.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00385; Chromo; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; RVT; 1.
DR SMART; SM00298; CHROMO; 1.
DR PROSITE; PS00598; CHROMO_1; 1.
DR PROSITE; PS00013; CHROMO_2; 1.
DR PROSITE; PS00994; INTEGRASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
KW Nuclear protein; Polyprotein; Protease.
FT NON_TER 1
SQ SEQUENCE 1243 AA; 143262 MW; C55999C5C80F58EE CRC64;
Query Match 100.0%; Score 34; DB 2; Length 1243;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db 336 TLREWL 341

RESULT 21
Q923T8_PSESX PRELIMINARY; PRT; 2066 AA.
AC Q923T8_PSESX PRELIMINARY;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Type I polyketide synthase.
GN Name=cfa7;
OS Pseudomonas syringae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=glycinea PG180;
RX MEDLINE=99080036; PubMed=9860992; DOI=10.1073/pnas.95.26.15469;
RA Rangasamy V., Jiralerspong S., Parry R., Bender C.L.;
RT "Biosynthesis of the Pseudomonas polyketide coronafacic acid requires
monofunctional and multifunctional polyketide synthase proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:15469-15474(1998).
DR EMBL; AF098795; AAD03048.1; -; Genomic_DNA.
DR HSP; P25715; 1MLA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR001227; ACP_transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR006163; Phosphateth bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF00550; PP-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
SQ SEQUENCE 2066 AA; 221022 MW; EEF3F9DC8D20F751 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 2066;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1316 TLREWL 1321

RESULT 22
Q87W69_PSESX PRELIMINARY; PRT; 2066 AA.
AC Q87W69_PSESX PRELIMINARY;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coronafacic acid polyketide synthetase II.
GN Name=cfa7; OrderedLocusName=PSPRO4687;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
```

```
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidse T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collmer A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
Pseudomonas syringae pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016853; AAO58131.1; -; Genomic_DNA.
DR HSP; P25715; 1MLA.
DR TIGR; PSPT04687; -.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR001227; ACP_transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR006163; Phosphateth bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF00550; PP-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
KW Complete proteome.
SQ SEQUENCE 2066 AA; 221220 MW; E9C6D0274BB901DF CRC64;

Query Match 100.0%; Score 34; DB 2; Length 2066;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1316 TLREWL 1321

RESULT 23
Q6JD66_TOXGO PRELIMINARY; PRT; 5072 AA.
AC Q6JD66_TOXGO PRELIMINARY;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Eukaryotic initiation factor-2 alpha kinase-A.
OS Tokoplasma gondii.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC Toxoplasma.
OX NCBI_TaxID=5811;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14989696;
RA Sullivan W.J. Jr., Narasimhan J., Bhatti M.M., Wek R.C.;
RT "Parsite-specific eIF2 (eukaryotic initiation factor-2) kinase
required for stress-induced translation control.";
RL Biochem. J. 380:523-531(2004).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AY518936; AAS48463.1; -; mRNA.
```



```

DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0003743; P:translation initiation factor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR008627; GETHR.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF05671; GETHR; 10.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Initiation factor; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 5072 AA; 541699 MW; 00FD7BED9848E58 CRC64;

Query Match 100.0%; Score 34; DB 2; Length 5072;
Best Local Similarity 100.0%; Pred. No. 4.9e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 4707 TLREWL 4712

RESULT 24
Q6FEL5 ACIAD
ID Q6FEL5 ACIAD PRELIMINARY; PRT; 81 AA.
AC Q6FEL5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative acyl carrier protein (ACP).
GN OrderedLocNames=ACIAD0569;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed15514110; DOI=10.1093/nar/gkh910;
RA Bathe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveillier S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
CC -!- SIMILARITY: Contains 1 acyl carrier domain.
DR EMBL; CR543861; CAG67493.1; -; Genomic DNA.
DR GO; GO:0000036; F:acyl carrier activity; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR InterPro; IPR003231; Acyl_carrier.
DR InterPro; IPR006163; Phosphateth_bind.
DR Pfam; PF00550; pp-binding; 1.
DR ProDom; PD000887; Acyl_carrier; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
KW Complete proteome; Phosphopantetheine.
SQ SEQUENCE 81 AA; 9345 MW; 58B296B82DF58F60 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 81;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 10 TLREWL 15

RESULT 25

```

```

Q70SL4_9MYRT
ID Q70SL4_9MYRT PRELIMINARY; PRT; 134 AA.
AC Q70SL4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Hauya heydeana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Hauya.
OX NCBI_TaxID=225332;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ550743; CAD79575.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR KW Isomerase.
FT NON TER 1
FT NON TER 134
SQ SEQUENCE 134 AA; 14294 MW; E2F33E5CBC99A21B CRC64;

Query Match 94.1%; Score 32; DB 2; Length 134;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 96 TLREWL 101

RESULT 26
Q70SM0_9MYRT
ID Q70SM0_9MYRT PRELIMINARY; PRT; 134 AA.
AC Q70SM0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Circaea cordata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Circaea.
OX NCBI_TaxID=13011;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ550737; CAD79569.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR KW Isomerase.
FT NON TER 1
FT NON TER 134
SQ SEQUENCE 134 AA; 14197 MW; 98EDD6882C1DB421 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 134;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6

```

```
Db 96 TLREWI 101
|||||
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 27
Q70SL7_9MYRT PRELIMINARY; PRT; 144 AA.
AC Q70SL7_2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=PgiC;
OS Fuchsia cylindracea.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Fuchsia.
OX NCBI_TaxID=225331;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ550740; CAD79572.1; -; Genomic DNA.
DR GO; GO:0004347; F:Glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; GGP_Isomerase.
DR Pfam; PF00342; GGP; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15248 MW; E487F612A63A4CPE CRC64;

Query Match 94.1%; Score 32; DB 2; Length 144;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 106 TLREWI 111
|||||

RESULT 28
Q656A5_ORYSA PRELIMINARY; PRT; 176 AA.
AC Q656A5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative terminal flower 1.
GN Name=P0596H06.13;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sakaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone P0596H06."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003620; BAD45362.1; -; Genomic_DNA.
DR GRAMENE; Q656A5;
DR InterPro; IPR001858; PBP.
DR InterPro; IPR008914; PEBP.
DR Pfam; PF01161; PBP; 1.
DR PRODOM; PD004330; PBP; 1.
DR PROSITE; PS01220; PBP; 1.
SQ SEQUENCE 176 AA; 19207 MW; F79EC00390487391 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 176;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 96 TLREWI 101
|||||
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 29
Q8D9L1_VIBVU PRELIMINARY; PRT; 195 AA.
AC Q8D9L1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=VV12586;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016805; AAO10937.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 195 AA; 21520 MW; 37B717876DF46C31 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 195;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 124 TLREWM 129
|||||

RESULT 30
Q6N1X5_RHOPA PRELIMINARY; PRT; 209 AA.
AC Q6N1X5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=RPA4277;
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
photosynthetic bacterium Rhodospseudomonas palustris."
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572606; CAE29718.1; -; Genomic DNA.
DR GO; GO:0007076; P:mitotic chromosome condensation; IEA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84DB040EE CRC64;

Query Match 94.1%; Score 32; DB 2; Length 209;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TLREWL 6
Db 109 TLREWL 114

RESULT 31
Q70SK1_9MYRT PRELIMINARY; PRT; 269 AA.
AC Q70SK1;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Lopezia grandiflora.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Lopezia.
OX NCBI_TaxID=225335;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ50756; CAD79588.1; -; Genomic DNA.
DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29125 MW; 0F67421013909157 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 269;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 96 TLREWL 101

RESULT 32
Q70SK3_9MYRT PRELIMINARY; PRT; 281 AA.
AC Q70SK3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiCb;
OS Megacora gracielanus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Megacora.
OX NCBI_TaxID=203949;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
```

```
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550754; CAD79586.1; -; Genomic DNA.
DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 281
SQ SEQUENCE 281 AA; 30422 MW; 68BB058315299731 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 281;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 96 TLREWL 101

RESULT 33
Q8RW74_CLAEP PRELIMINARY; PRT; 307 AA.
AC Q8RW74;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC2;
OS Clarkia epilobioides (Willow-herb clarkia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=145959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LDS 9605;
RX MEDLINE=22033959; PubMed=12038528;
RA Ford V.S., Gottlieb L.D.;
RT "Single mutations silence pgiC2 genes in two very recent
allotetraploid species of Clarkia.";
RL Evolution 56:699-707(2002).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ311754; CAC84516.1; -; Genomic DNA.
DR HSP; P08059; IGZD.
DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 307
SQ SEQUENCE 307 AA; 33553 MW; A62C334EBF309B25 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 307;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 1 TLREWL 6
```

Db 156 TLREWI 161

RESULT 34
TOC34_ARATH STANDARD; PRT; 313 AA.
AC Q38906; Q8L9G5; Q9PF74; Q9GDD2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Translocase of chloroplast 34 (34 kDa chloroplast outer envelope protein) (Grp-binding protein OEP34) (AtToc34).
GN Name=oc34; Synonyms=ocp34; OrderedLocusNames=At5g05000;
GN ORFNames=MUG13.14;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia;
RA Chen L.-J., Li H.-M.;
RT "Arabidopsis homologue of OEP34, a component of the pea chloroplast protein import apparatus."
RL (er) Plant Gene Register PGR95-140.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Columbia;
RX MEDLINE=20452907; PubMed=10998188;
RA Gutensohn M., Schulz B.I., Nicolay P., Fluegge U.I.;
RT "Functional analysis of the two Arabidopsis homologues of Toc34, a component of the chloroplast protein import apparatus."
RL Plant J. 27:771-783 (2000).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M., Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones."
RL DNA Res. 4:215-230 (1997).
[4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R., Feldmann K.A.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[5]
RN SUBCELLULAR LOCATION.
RX PubMed=12775770; DOI=10.1074/mcp.M300030-MCP200;
RA Ferro M., Salvi D., Brugiere S., Mitras S., Kowalski S., Louwagie M., Garin J., Joyard J., Rolland N.;
RT "Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana."
RL Mol. Cell. Proteomics 2:325-345 (2003).
CC -!- FUNCTION: Involved in protein import into chloroplasts. Seems to recognize chloroplast-designed precursor proteins and regulate their presentation to the translocation channel through GTP hydrolysis.
CC -!- SUBUNIT: Interacts with Toc75.
CC -!- SUBCELLULAR LOCATION: Chloroplast; outer membrane.
CC -!- SIMILARITY: Belongs to the TOC34 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC
CC EMBL; U43377; AAD09203.1; -; mRNA.

DR EMBL; AJ132696; CAC17699.1; -; mRNA.
DR EMBL; AB005245; BAB11522.1; -; Genomic_DNA.
DR EMBL; AY084447; AAM65983.1; -; mRNA.
DR InterPro; IPR006703; AIG1.
DR InterPro; IPR005688; Toc34.
DR Pfam; PF04548; AIG1; 1.
DR TIGRFAMs; TIGR00991; 3a0901s02IAP34; 1.
KW Chloroplast; GTP-binding; Membrane; Nucleotide-binding;
KW Outer membrane; Protein transport; Transport.
FT NP_BIND 45 52 GTP (Potential).
FT NP_BIND 92 96 GTP (Potential).
FT NP_BIND 154 157 M -> I (in Ref. 4).
FT CONFLICT 44 44 G -> S (in Ref. 4).
FT CONFLICT 47 47 L -> P (in Ref. 3 and 4).
FT CONFLICT 299 299 L -> P (in Ref. 3 and 4).
SQ SEQUENCE 313 AA; 34723 MW; 1C90E435B2EE1F2D CRC64;
Query Match 94.1%; Score 32; DB 1; Length 313;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TLREWL 6
Db 6 TLREWI 11
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]

```

OX NCBI_TaxID=264201;
RN [1]
RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA PubMed=15073324;
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA Fartmann B., Brandt P., Nyakatura G.J., Drooge M., Frishman D.,
RA Rattai T., Mewes H.-W., Wagner M.;
RT "Illuminating the evolutionary history of chlamydiae.";
EL Science 304:728-730(2004).
DR EMBL: BX908798; CAF24207.1; -; Genomic_DNA.
DR InterPro: IPR000644; CBS.
DR InterPro: IPR002550; DUF21.
DR Pfam: PF00571; CBS; 1.
DR Pfam: PF01595; DUF21; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 342 AA; 38301 MW; D10713478B85A64E CRC64;

Query Match 94.1%; Score 32; DB 2; Length 342;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TLREWL 6
DB 31 TLREWI 36
|||||

RESULT 37
O22092 9LILI PRELIMINARY; PRT; 357 AA.
AC O22092;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea nipponica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64700;
RN [1]
RX MEDLINE=98172211; PubMed=9511221;
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL: AB006617; BAA22035.1; -; mRNA.
DR HSP; P06744; IAT.
DR GO: GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
DR GO: GO:0016853; F:isomerase activity; IEA.
DR GO: GO:0006094; P:gluconeogenesis; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000886; ER target S.
DR InterPro: IPR001672; G6P_Isomerase.
DR Pfam: PF00342; PGI; 1.
DR PRINTS: PR00662; G6PISOMERASE.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
SQ SEQUENCE 357 AA; 39037 MW; 4E2CEFEED30FA3AE9 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TLREWL 6
DB 153 TLREWI 158
|||||

RESULT 39
O22094 9LILI PRELIMINARY; PRT; 357 AA.
AC O22094;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea septemloba.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64701;
RN [1]
RX MEDLINE=98172211; PubMed=9511221;

```

```

OY 1 TLREWL 6
DB 153 TLREWI 158
|||||

RESULT 38
O22093 9LILI PRELIMINARY; PRT; 357 AA.
AC O22093;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea quinqueloba.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64477;
RN [1]
RX MEDLINE=98172211; PubMed=9511221;
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262(1997).
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL: AB006618; BAA22036.1; -; mRNA.
DR HSP; P06744; IAT.
DR GO: GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
DR GO: GO:0016853; F:isomerase activity; IEA.
DR GO: GO:0006094; P:gluconeogenesis; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR000886; ER target S.
DR InterPro: IPR001672; G6P_Isomerase.
DR Pfam: PF00342; PGI; 1.
DR PRINTS: PR00662; G6PISOMERASE.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 357
SQ SEQUENCE 357 AA; 39037 MW; 4E2CEFEED30FA3AE9 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TLREWL 6
DB 153 TLREWI 158
|||||

RESULT 39
O22094 9LILI PRELIMINARY; PRT; 357 AA.
AC O22094;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea septemloba.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OX NCBI_TaxID=64701;
RN [1]
RX MEDLINE=98172211; PubMed=9511221;

```

```
RA Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262 (1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC EMBL: AB006619; BAA22037.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER target S.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00562; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1 357
FT SEQUENCE 357 AA; 39029 MW; 046D30C542200FF4 CRC64;
SQ

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 153 TLREWI 158

RESULT 40
O22095 9LILI
ID O22095_9LILI PRELIMINARY; PRT; 357 AA.
AC O22095;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea tenuipes.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OC NCBI_TaxID=64476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262 (1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006620; BAA22038.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER target S.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00562; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1 357
FT SEQUENCE 357 AA; 39029 MW; 046D30C542200FF4 CRC64;
SQ
```

```
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1 357
FT SEQUENCE 357 AA; 39078 MW; DE9C1AE79030E6AD CRC64;
SQ

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 153 TLREWI 158

RESULT 41
O24644 9LILI
ID O24644_9LILI PRELIMINARY; PRT; 357 AA.
AC O24644;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
OS Dioscorea gracillima.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Dioscoreales; Dioscoreaceae;
OC Dioscorea.
OC NCBI_TaxID=64699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Kawabe A., Miyashita N.T., Terauchi R.;
RT "Phylogenetic relationship among the section Stenophora in the genus
RT Dioscorea based on the analysis of nucleotide sequence variation in
RT the phosphoglucose isomerase (Pgi) locus.";
RL Genes Genet. Syst. 72:253-262 (1997).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AB006616; BAA22034.1; -; mRNA.
DR EMBL; AB006615; BAA22033.1; -; mRNA.
DR HSSP; P06744; 1IAT.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000886; ER target S.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00562; G6PISOMERASE.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1 357
FT SEQUENCE 357 AA; 39021 MW; 08DC40108A3B165B CRC64;
SQ

Query Match 94.1%; Score 32; DB 2; Length 357;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 153 TLREWI 158

RESULT 42
O4SIN4_TETNG
ID O4SIN4_TETNG PRELIMINARY; PRT; 368 AA.
AC O4SIN4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
```

```
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 21 SCAF14577, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00017591001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Telostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet P., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaut S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bismont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBSJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; CA901014577; CAF99498.1; -; Genomic DNA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW NON TER 368 368
SQ SEQUENCE 368 AA; 40576 MW; 7C860ECA3DDF1BE9 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 368;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
DB 351 TLREWM 356
|||||:

RESULT 43
Q70SK2 9MYRT PRELIMINARY; PRT; 371 AA.
AC Q70SK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Hauya elegans.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Hauya.
OX NCBI_TaxID=13081;
RN [1]
RP NUCLEOTIDE SEQUENCE.

Query Match 94.1%; Score 32; DB 2; Length 371;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
DB 351 TLREWM 356
|||||:

RESULT 44
Q70SK7 9MYRT PRELIMINARY; PRT; 371 AA.
AC Q70SK7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Epilobium canum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Epilobium.
OX NCBI_TaxID=33131;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550750; CAD79582.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40544 MW; 51994754C9F26FBC CRC64;

Query Match 94.1%; Score 32; DB 2; Length 371;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
DB 96 TLREWI 101
|||||:
```

```
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550755; CAD79587.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40579 MW; A30A540BAC1DE9FB CRC64;

Query Match 94.1%; Score 32; DB 2; Length 371;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
DB 96 TLREWI 101
|||||:

RESULT 44
Q70SK7 9MYRT PRELIMINARY; PRT; 371 AA.
AC Q70SK7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=pgiC;
OS Epilobium canum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Epilobium.
OX NCBI_TaxID=33131;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ550750; CAD79582.1; -; Genomic DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40544 MW; 51994754C9F26FBC CRC64;

Query Match 94.1%; Score 32; DB 2; Length 371;
Best Local Similarity 83.3%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TLREWL 6
DB 96 TLREWI 101
|||||:
```

```
Db          96 TLREWI 101

RESULT 45
Q70SK5_9MYRT
ID Q70SK5_9MYRT PRELIMINARY; PRT; 373 AA.
AC Q70SK5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=PgiC2;
OS Epilobium brachycarpum.
OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Epilobium.
OX NCBI_TaxID=29754;
RN [1]
RP NUCLEOTIDE SEQUENCE
RA Gottlieb L.D., Ford V.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ350752; CAD79584.1; -; Genomic_DNA.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 1
FT NON_TER 373
SQ SEQUENCE 373 AA; 40728 MW; 2425C5F1F6D1B4FE CRC64;

Query Match 94.1%; Score 32; DB 2; Length 373;
Best Local Similarity 83.3%; Pred. No. 7.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 96 TLREWI 101

RESULT 46
Q8L6V2_9MYRT
ID Q8L6V2_9MYRT PRELIMINARY; PRT; 380 AA.
AC Q8L6V2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=PgiC2b;
OS Clarkia dudleyana.
OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=188038;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ford V.S., Gottlieb L.D.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ343727; CAD24790.1; -; Genomic_DNA.
DR HSSP; P08059; IGZD.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR NUCLEOTIDE SEQUENCE.
RA STRAIN=LDG 8624a;
RP STRAIN=LDG 8624a;
RX MEDLINE=22189021; PubMed=12200488;
RA "The 5' leader of plant PgiC has an intron; the leader shows both the
RT loss and maintenance of constraints compared with introns and exons in
RT the coding region.";
RL Mol. Biol. Evol. 19:1613-1623(2002).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ343727; CAD24790.1; -; Genomic_DNA.
DR HSSP; P08059; IGZD.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR STRAIN=LDG 8624a;
```

```
RX MEDLINE=22189021; PubMed=12200488;
RA Gottlieb L.D., Ford V.S.;
RT "The 5' leader of plant PgiC has an intron; the leader shows both the
RT loss and maintenance of constraints compared with introns and exons in
RT the coding region.";
RL Mol. Biol. Evol. 19:1613-1623(2002).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ437278; CAD24791.1; -; Genomic_DNA.
DR HSSP; P08059; IGZD.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Gluconeogenesis; Glycolysis; Isomerase.
FT NON_TER 380
FT NON_TER 380
SQ SEQUENCE 380 AA; 41687 MW; 6871252AFC5B066 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 380;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
Db 226 TLREWI 231

RESULT 47
Q8L6V3_9MYRT
ID Q8L6V3_9MYRT PRELIMINARY; PRT; 380 AA.
AC Q8L6V3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
GN Name=PgiC2a;
OS Clarkia dudleyana.
OC Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; Myrtales; Onagraceae; Clarkia.
OX NCBI_TaxID=188038;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ford V.S., Gottlieb L.D.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
phosphate.
CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the GPI family.
DR EMBL; AJ437277; CAD24790.1; -; Genomic_DNA.
DR HSSP; P08059; IGZD.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR STRAIN=LDG 8624a;
```


Mon May 15 11:35:31 2006

DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON TER 380 380
 SQ SEQUENCE 380 AA; 41631 MW; AC59DFE13021CE7F CRC64;
 Query Match 94.1%; Score 32; DB 2; Length 380;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 226 TLREWI 231
 RESULT 48
 ID Q8L6V4_9MYRT PRELIMINARY; PRT; 380 AA.
 AC Q8L6V4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
 GN Name=pgiC;
 OS Clarkia dudleyana.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; Myrtales; Onagraceae; Clarkia.
 OX NCBI_TaxID=188038;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LDG 8624a;
 RA Ford V.S., Gottlieb L.D.;
 RT "Reassessment of phylogenetic relationships in Clarkia sect.
 Symphoricarica.";
 RL Am. J. Bot. 90:284-292(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LDG 8624a;
 RA Gottlieb L.D., Ford V.S.;
 RT "The 5' leader of plant PgiC has an intron: the leader shows both the
 loss and maintenance of constraints compared with introns and exons in
 the coding region.";
 RL Mol. Biol. Evol. 19:1613-1623(2002).
 CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
 phosphate.
 CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the GPI family.
 DR EMBL; AJ437276; CAD24789.1; -; Genomic_DNA.
 DR HSP; Q9N182; IHMS.
 DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; P:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON TER 380 380
 SQ SEQUENCE 380 AA; 41659 MW; F6E18DF30212F586 CRC64;
 Query Match 94.1%; Score 32; DB 2; Length 380;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6

DB 226 TLREWI 231
 RESULT 49
 ID Q8RW61_CLAUN PRELIMINARY; PRT; 380 AA.
 AC Q8RW61;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
 GN Name=pgiC;
 OS Clarkia unguiculata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; Myrtales; Onagraceae; Clarkia.
 OX NCBI_TaxID=3337;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LDG-8829-8d-3A;
 RX MEDLINE=22033959; PubMed=12038528;
 RA Ford V.S., Gottlieb L.D.;
 RT "Single mutations silence PgiC2 genes in two very recent
 allotetraploid species of Clarkia.";
 RL Evolution 56:699-707(2002).
 CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-
 phosphate.
 CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the GPI family.
 DR EMBL; AJ312370; CAC86124.1; -; Genomic_DNA.
 DR HSP; Q9N1E2; IHMS.
 DR GO; GO:0004347; P:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; P:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON TER 380 380
 SQ SEQUENCE 380 AA; 41762 MW; CCC72616F23BFB7F CRC64;
 Query Match 94.1%; Score 32; DB 2; Length 380;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TLREWL 6
 DB 226 TLREWI 231
 RESULT 50
 ID Q7OSL1_LUDPE PRELIMINARY; PRT; 381 AA.
 AC Q7OSL1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Cytosolic phosphoglucose isomerase (EC 5.3.1.9) (Fragment).
 GN Name=pgiC;
 OS Ludwigia peploides (Floating primrose-willow).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; Myrtales; Onagraceae; Ludwigia.
 OX NCBI_TaxID=13119;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Gottlieb L.D., Ford V.S.;
 RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate = D-fructose 6-

CC phosphate.
 CC -1- PATHWAY: Involved in glycolysis and in gluconeogenesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the GPI family.
 DR EMBL; AJ550746; CAD79578.1; -; Genomic DNA.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_isomerase.
 DR Pfam; PF00342; FGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P GLUCOSE ISOMERASE 1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON_TER 1
 FT NON_TER 381
 SQ SEQUENCE 381 AA; 41846 MW; F54CC9FC3161EDC2 CRC64;

Query Match 94.1%; Score 32; DB 2; Length 381;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLREWL 6
 Db 106 TLREWI 111

Search completed: May 12, 2006, 10:51:28
 Job time : 74.4615 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 72.6496 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-35
Perfect score: 56
Sequence: 1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq 21:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*
9: geneseqp2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	10	AAV22355	Aay22355 TPO recep
2	56	100.0	10	AAAB16979	Abb16979 TPO-mimet
3	56	100.0	10	ABB72865	Abb72865 TPO mimet
4	56	100.0	10	ADJ73016	Adj73016 TPO mimet
5	56	100.0	10	ADJ52651	Adj52651 CH1 delet
6	56	100.0	10	ADJ51612	Adj51612 CH1 delet
7	41	73.2	780	ADN20123	Adn20123 Bacterial
8	41	73.2	787	ADA36768	Ada36768 Acinetoba
9	40	71.4	96	AAW92258	Aam92258 Human dig
10	40	71.4	353	AAW91714	Abm91714 M. xanthu
11	40	71.4	363	ADN21256	Adn21256 Bacterial
12	39	69.6	95	AAAB54291	Aab54291 Human pan
13	39	69.6	558	AAAB7456	Aab7456 D-aminOac
14	38	67.9	358	ABO62198	AbO62198 Klebsiell
15	38	67.9	536	AAW64365	Aaw64365 Mycobacte
16	38	67.9	536	AAW81732	Aaw81732 M. tuberc
17	38	67.9	536	AAV39019	Aay39019 M. tuberc
18	38	67.9	536	AAV39162	Aay39162 M. tuberc
19	37	66.1	103	AAU54158	Aau54158 Propionib
20	37	66.1	103	ABMS0677	Abms0677 Pseudomon
21	37	66.1	206	ABO80351	AbO80351 Pseudomon
22	37	66.1	271	ABO84669	AbO84669 Mouse can
23	37	66.1	297	AAAG2549	Aag2549 C glutami
24	37	66.1	423	ADT57937	Adt57937 Plant pol

25	37	66.1	605	4	AAG91159	Aag91159 C glutami
26	37	66.1	746	8	ADS29923	Ads29923 Bacterial
27	36	64.3	66	4	AAU64755	Aau64755 Propionib
28	36	64.3	66	6	ABM61274	Abm61274 Human ORF
29	36	64.3	159	3	AAB41784	Aab41784 Human ORF
30	36	64.3	290	9	ABM97186	Abm97186 M. xanthu
31	36	64.3	336	5	ABB84757	Abb84757 DNA polym
32	36	64.3	336	5	ABB84815	Abb84815 DNA polym
33	36	64.3	343	4	ABG01473	Abg01473 Novel hum
34	36	64.3	355	6	ABU41540	Abu41540 Protein e
35	36	64.3	385	8	ADH18901	Adh18901 Human cel
36	36	64.3	387	7	ABO81727	AbO81727 Pseudomon
37	36	64.3	415	6	ABM15899	Abm15899 Mycobacte
38	36	64.3	423	5	ABB89453	Abb89453 Human pol
39	36	64.3	436	7	RAO30830	Rao30830 Human cel
40	36	64.3	479	5	ABN90892	Abn90892 Herbicida
41	36	64.3	479	8	ADT56368	Adt56368 Plant pol
42	36	64.3	479	9	ADV42198	Adv42198 Glycosylt
43	36	64.3	517	3	AAAB41790	Aab41790 Human ORF
44	36	64.3	517	4	AAAM93202	Aam93202 Human pol
45	36	64.3	517	4	AAAM93206	Aam93206 Human pol
46	36	64.3	517	8	ADL30558	Adl30558 Human pro
47	36	64.3	517	8	ADL30566	Adl30566 Human pro
48	36	64.3	527	8	ADH18902	Adh18902 Human cel
49	36	64.3	560	8	ADQ21291	Adq21291 Human sef
50	36	64.3	578	4	AAAG81257	Aag81257 Human AFP
51	36	64.3	579	2	AAAY30812	Aay30812 Human sec
52	36	64.3	594	6	ABU11815	Abu11815 Human MDD
53	36	64.3	594	6	ABU11514	Abu11514 Human MDD
54	36	64.3	595	3	AAAB12121	Aab12121 Hydrophob
55	36	64.3	662	5	ABP42016	Abp42016 Human ova
56	36	64.3	662	4	ABG22092	Abg22092 Novel hum
57	36	64.3	1021	7	ABO70459	AbO70459 Pseudomon
58	35	62.5	10	2	AAV22356	Aay22356 TPO recep
59	35	62.5	10	3	AAAB16980	Abb16980 TPO-mimet
60	35	62.5	10	5	ABB72866	Abb72866 TPO mimet
61	35	62.5	10	7	ADJ73017	Adj73017 TPO mimet
62	35	62.5	10	8	ADJ52652	Adj52652 CH1 delet
63	35	62.5	10	8	ADJ51613	Adj51613 CH1 delet
64	35	62.5	11	2	AAV22372	Aay22372 TPO recep
65	35	62.5	15	2	AAV22403	Aay22403 TPO recep
66	35	62.5	15	2	AAV22375	Aay22375 TPO recep
67	35	62.5	15	2	AAV22402	Aay22402 TPO recep
68	35	62.5	20	7	ADC99211	Adc99211 Cancer-re
69	35	62.5	36	6	ADA04633	Ada04633 IGF-IR bi
70	35	62.5	36	7	ADH95846	Adh95846 Insulin g
71	35	62.5	36	8	ADL68537	Adl68537 IGF-IR/IR
72	35	62.5	36	8	ADM38382	Adm38382 Insulin a
73	35	62.5	50	8	ADK01720	Adk01720 Hepatitis
74	35	62.5	113	6	ABU49297	Abu49297 Protein e
75	35	62.5	140	5	ADK35767	Adk35767 Novel hum
76	35	62.5	167	3	AAV75214	Aay75214 Neisseria
77	35	62.5	169	3	AAV75215	Aay75215 Neisseria
78	35	62.5	217	7	ABO79586	AbO79586 Pseudomon
79	35	62.5	220	3	AAV75216	Aay75216 Neisseria
80	35	62.5	261	8	ADJ35064	Adj35064 Xylanase
81	35	62.5	264	6	ABU41805	Abu41805 Protein e
82	35	62.5	284	4	ABE66010	AbE66010 Drosophil
83	35	62.5	296	5	ABG98039	Abg98039 Mutant ha
84	35	62.5	370	6	ABU15760	Abu15760 Protein e
85	35	62.5	371	7	ABO77020	AbO77020 Pseudomon
86	35	62.5	396	8	ADT60239	Adt60239 Plant pol
87	35	62.5	437	9	ABR41989	AbR41989 Coniochae
88	35	62.5	437	9	ADU97707	Adu97707 Coniochae
89	35	62.5	444	5	AAO22151	Aao22151 Ramoplani
90	35	62.5	447	3	AAAG29319	Aag29319 Arabidops
91	35	62.5	452	8	ADS64781	Ads64781 Arthriniu
92	35	62.5	482	3	AAAG29318	Aag29318 Arabidops
93	35	62.5	482	5	ABN92055	Abn92055 Herbicida
94	35	62.5	482	7	ADB91974	Adb91974 Acyltrans
95	35	62.5	490	7	ADB70195	Adb70195 C. neofor
96	35	62.5	492	7	ABM86990	Abm86990 Rice abio
97	35	62.5	495	4	AAAB94358	Aab94358 Human pro

98	35	62.5	500	6	ABP57627	Abp57627 S. muraya	171	34	60.7	230	8	ADx87466	Adx87466 Plant ful
99	35	62.5	507	8	ADT60952	Adt60952 Plant pol	172	34	60.7	236	4	AAU63835	Aau63835 Propionib
100	35	62.5	520	4	ABU52637	Abu52637 Human tra	173	34	60.7	236	4	ABM60354	Abm60354 Propionib
101	35	62.5	520	4	ABU52637	Abu52637 Human tra	173	34	60.7	236	4	ABM60354	Abm60354 Propionib
102	35	62.5	520	8	ADS23199	Ads23199 Bacterial	175	34	60.7	264	8	ADx66559	Adx66559 Plant ful
103	35	62.5	547	7	ABO77986	AbO77986 Pseudomon	176	34	60.7	293	9	ADN21928	Adn21928 Bacterial
104	35	62.5	607	5	AAE22155	Aae22155 Human TRN	177	34	60.7	296	7	ABM96423	Abm96423 M. xanthu
105	35	62.5	607	5	ADA54308	Ada54308 Human pro	178	34	60.7	303	7	ABM87579	Abm87579 Rice abio
106	35	62.5	607	7	ADC53389	Adc53389 Human N-a	179	34	60.7	318	7	ABO78573	AbO78573 Pseudomon
107	35	62.5	607	7	ADE47686	Ade47686 Human NOV	180	34	60.7	343	7	ADD43962	Add43962 Chlamydia
108	35	62.5	607	7	ADE47688	Ade47688 Human NOV	181	34	60.7	343	9	AEA19254	Aea19254 Chlamydia
109	35	62.5	607	7	ADE47690	Ade47690 Human NOV	182	34	60.7	345	8	ADN80524	Adn80524 Plant ful
110	35	62.5	607	8	ADJ78958	Adj78958 Human NOV	183	34	60.7	351	8	ADN21155	Adn21155 Bacterial
111	35	62.5	607	8	ADJ78956	Adj78956 Human NOV	184	34	60.7	375	2	AAW20731	Aaw20731 H. pylori
112	35	62.5	607	8	ADJ78960	Adj78960 Human NOV	185	34	60.7	380	4	ABB06901	Abb06901 Micromono
113	35	62.5	620	7	ABO72087	AbO72087 Pseudomon	186	34	60.7	380	6	ABP99317	Abp99317 Orthosomy
114	35	62.5	944	7	ADE08076	Ade08076 Novel pro	187	34	60.7	380	8	ADT59120	Adt59120 Plant pol
115	35	62.5	944	9	ADU40446	Adu40446 Novel hum	188	34	60.7	413	8	ADx68804	Adx68804 Plant ful
116	35	62.5	948	4	ABB63850	Abb63850 Drosophil	189	34	60.7	437	2	AAy37663	Aay37663 Amino aci
117	35	62.5	971	4	ABG24290	Abg24290 Novel hum	190	34	60.7	444	8	ADx72887	Adx72887 Plant ful
118	35	62.5	1054	9	ABE69202	Abe69202 Human mod	191	34	60.7	454	6	ABU45271	Abu45271 Protein e
119	35	62.5	1117	8	ADM64308	Adm64308 High-temp	192	34	60.7	457	6	ABU28329	Abu28329 Protein e
120	35	62.5	1234	6	ABM68967	Abm68967 Phototrab	193	34	60.7	457	6	ABU32229	Abu32229 Protein e
121	35	62.5	1515	6	ABU19676	Abu19676 Protein e	194	34	60.7	457	6	ABU15420	Abu15420 Protein e
122	35	62.5	2600	6	ABU22788	Abu22788 Protein e	195	34	60.7	457	6	ABU50197	Abu50197 Protein e
123	34.5	61.6	275	3	AAy87221	Aay87221 Human sec	196	34	60.7	457	8	ADS45219	Ads45219 Bacterial
124	34.5	61.6	275	4	AAE06203	Aae06203 Human gen	197	34	60.7	458	6	ABU48015	Abu48015 Protein e
125	34.5	61.6	275	5	ABG34027	Abg34027 Human sec	198	34	60.7	458	6	ABU47154	Abu47154 Protein e
126	34.5	61.6	319	3	AAy87223	Aay87223 Human sec	199	34	60.7	461	7	ABO65735	AbO65735 Klebsiell
127	34.5	61.6	319	4	AAE06205	Aae06205 Human gen	200	34	60.7	488	6	ABU30709	Abu30709 Protein e
128	34.5	61.6	319	4	AAE06132	Aae06132 Human gen	201	34	60.7	490	5	ABP65879	Abp65879 Bifidobac
129	34.5	61.6	319	5	ABG34029	Abg34029 Human sec	202	34	60.7	548	2	AAR22197	Aar22197 Sequence
130	34.5	61.6	319	5	ABG33955	Abg33955 Human sec	203	34	60.7	575	8	ADY04932	Ady04932 Plant ful
131	34.5	61.6	319	6	ADA57361	Ada57361 Human sec	204	34	60.7	587	7	ABM90283	Abm90283 Rice abio
132	34.5	61.6	319	6	ADA41242	Ada41242 Human sec	205	34	60.7	609	7	ABO83801	AbO83801 Pseudomon
133	34.5	61.6	319	6	ABR48003	AbR48003 Human sec	206	34	60.7	660	2	AAR70013	Aar70013 Tick carb
134	34.5	61.6	320	3	AAy87155	Aay87155 Human sec	207	34	60.7	801	8	ADx88989	Adx88989 Plant ful
135	34.5	61.6	556	7	ADD19311	AdD19311 Human sec	208	34	60.7	859	8	ADQ37037	Adq37037 Cell prol
136	34	60.7	9	7	ADJ73026	Adj73026 TPO mimet	209	34	60.7	927	7	ABO64572	AbO64572 Klebsiell
137	34	60.7	9	8	ADJ51622	Adj51622 CHI delet	210	34	60.7	977	3	AAy51199	Aay51199 Arabidops
138	34	60.7	11	2	AAy22367	Aay22367 TPO recep	211	34	60.7	992	3	AAy51198	Aay51198 Arabidops
139	34	60.7	11	2	AAy22365	Aay22365 TPO recep	212	34	60.7	1051	4	ABB60157	Abb60157 Drosophil
140	34	60.7	11	2	AAy22370	Aay22370 TPO recep	213	34	60.7	1115	3	AAy51197	Aay51197 Arabidops
141	34	60.7	11	3	AAy16989	Aay16989 TPO-mimet	214	34	60.7	1115	5	AAU81522	Aau81522 C.elegans
142	34	60.7	11	3	AAy16986	Aay16986 TPO-mimet	215	34	60.7	1214	7	ABP96831	Abp96831 Pucative
143	34	60.7	11	3	AAy16988	Aay16988 TPO-mimet	216	34	60.7	1221	7	ADM25443	Adm25443 Hyperther
144	34	60.7	11	5	ABB72872	Abb72872 TPO mimet	217	34	60.7	1529	3	AAy96744	Aay96744 A. terreu
145	34	60.7	11	5	ABB72875	Abb72875 TPO mimet	218	34	60.7	1711	8	ADN46550	Adn46550 Thermococ
146	34	60.7	11	5	ABB72874	Abb72874 TPO mimet	219	34	60.7	4437	6	ABU11384	Abu11384 Protein e
147	34	60.7	11	7	ADJ73023	Adj73023 TPO mimet	220	34	60.7	19938	6	ABP76680	Abp76680 Streptomy
148	34	60.7	11	7	ADJ73025	Adj73025 TPO mimet	221	33.5	59.8	633	9	ABM96728	Abm96728 M. xanthu
149	34	60.7	11	8	ADJ52658	Adj52658 CHI delet	222	33	58.9	7	5	AAO18380	Aao18380 Lead stru
150	34	60.7	11	8	ADJ52661	Adj52661 CHI delet	223	33	58.9	15	2	AAy22406	Aay22406 TPO recep
151	34	60.7	11	8	ADJ52660	Adj52660 CHI delet	224	33	58.9	54	9	ABE18127	AbE18127 Lactobaci
152	34	60.7	11	8	ADJ51619	Adj51619 CHI delet	225	33	58.9	57	4	AAU57120	Aau57120 Propionib
153	34	60.7	11	8	ADJ51621	Adj51621 CHI delet	226	33	58.9	57	6	ABM53639	Abm53639 Propionib
154	34	60.7	13	2	AAy22407	Aay22407 TPO recep	227	33	58.9	85	8	ADY25087	Ady25087 Plant ful
155	34	60.7	50	2	AAy07987	Aay07987 Human sec	228	33	58.9	103	4	AAU43490	Aau43490 Propionib
156	34	60.7	63	5	ABP00760	Abp00760 Human ORF	229	33	58.9	103	6	ABM40009	Abm40009 Propionib
157	34	60.7	73	4	AAy92259	Aay92259 Human dig	230	33	58.9	122	4	AAU59364	Aau59364 Propionib
158	34	60.7	76	2	AAy74245	Aay74245 Human pro	231	33	58.9	122	6	ABM55883	Abm55883 Propionib
159	34	60.7	80	4	AAy82693	Aay82693 Human imm	232	33	58.9	134	4	AAU52483	Aau52483 Propionib
160	34	60.7	124	7	ABO65668	AbO65668 Klebsiell	233	33	58.9	134	6	ABM49002	Abm49002 Propionib
161	34	60.7	148	7	ABO84012	AbO84012 Pseudomon	234	33	58.9	135	7	ABO68543	AbO68543 Pseudomon
162	34	60.7	155	7	ABO81847	AbO81847 Pseudomon	235	33	58.9	137	6	AAU42999	Aau42999 Propionib
163	34	60.7	157	7	ABO63703	AbO63703 Klebsiell	236	33	58.9	137	6	ABM39518	Abm39518 Propionib
164	34	60.7	193	9	ADY18608	Ady18608 PRO polyp	237	33	58.9	144	8	ADY23540	Ady23540 Plant ful
165	34	60.7	196	8	ADL08563	AdL08563 Human tum	238	33	58.9	156	4	ABG03602	Abg03602 Novel hum
166	34	60.7	202	7	ABE64982	AbE64982 Human pro	239	33	58.9	160	6	AAU44716	Aau44716 Propionib
167	34	60.7	213	4	AAy19897	Aay19897 Polyketid	240	33	58.9	160	6	ABM41235	Abm41235 Propionib
168	34	60.7	215	6	ADA34477	Ada34477 Acinetoba	241	33	58.9	164	7	ABO70547	AbO70547 Pseudomon
169	34	60.7	219	7	ABO67958	AbO67958 Pseudomon	242	33	58.9	179	4	AAU36304	Aau36304 Pseudomon
170	34	60.7	228	8	ADx79685	Adx79685 Plant ful	243	33	58.9	179	6	ABU38519	Abu38519 Protein e

244	33	58.9	196	8	AD013837	Ad013837 Protein e	317	33	58.9	693	4	ABG28564	Abg28564 Novel hum
245	33	58.9	196	9	AB37723	Aeb37723 L. pneumo	318	33	58.9	702	4	ABG27190	Abg27190 Human hum
246	33	58.9	196	9	AB41033	Aeb41033 L. pneumo	319	33	58.9	702	4	ADC33307	Adc33307 Human nov
247	33	58.9	199	8	ADM90889	Adm90889 Human pha	320	33	58.9	713	8	ADS21802	Ads21802 Bacterial
248	33	58.9	205	8	ADP30231	Adp30231 Human sec	321	33	58.9	754	8	ADS30847	Ads30847 Bacterial
249	33	58.9	205	8	ADP30232	Adp30232 Human sec	322	33	58.9	774	4	AAU36453	Aau36453 Pseudomon
250	33	58.9	210	7	ADF58756	Adf58756 Human pol	323	33	58.9	774	6	ABU38785	Abu38785 Protein e
251	33	58.9	218	6	ABU40896	Abu40896 Protein e	324	33	58.9	777	7	ABU74951	Abu74951 Pseudomon
252	33	58.9	218	8	ADN26233	Adn26233 Bacterial	325	33	58.9	790	7	ABM86794	Abm86794 Rice abio
253	33	58.9	222	7	ADF07151	Adf07151 Bacterial	326	33	58.9	802	2	AAU56550	Aau56550 Cold accl
254	33	58.9	232	8	ADN47634	Adn47634 Thermococ	327	33	58.9	826	6	ABU28255	Abu28255 Protein e
255	33	58.9	249	7	ADC01166	Adc01166 Enterohae	328	33	58.9	828	2	AAU04357	Aau04357 E. coli p
256	33	58.9	252	4	ABG28561	Abg28561 Novel hum	329	33	58.9	850	2	AAU86955	Aau86955 E. coli p
257	33	58.9	256	8	ADK94555	Adk94555 Plant ful	330	33	58.9	858	6	ABU47698	Abu47698 Protein e
258	33	58.9	257	8	ADN26848	Adn26848 Bacterial	331	33	58.9	858	6	ABU15100	Abu15100 Protein e
259	33	58.9	257	8	ADN26596	Adn26596 Bacterial	332	33	58.9	858	6	ABU47261	Abu47261 Protein e
260	33	58.9	258	8	ADN62549	Adn62549 A. thalia	333	33	58.9	860	6	ABU32346	Abu32346 Protein e
261	33	58.9	261	2	AAW88462	Aaw88462 Bacillus	334	33	58.9	863	7	ABO64567	Abu64567 Klebsiell
262	33	58.9	261	8	ADT58437	Adt58437 Plant pol	335	33	58.9	901	9	AEB17427	Aeb17427 E. coli K
263	33	58.9	267	8	ADT58437	Adt58437 Plant pol	336	33	58.9	901	9	AEB17427	Aeb17427 E. coli K
264	33	58.9	272	4	AAH81225	Aah81225 Mycobacte	337	33	58.9	923	8	ADT04046	Adt04046 Human pro
265	33	58.9	272	5	AAU91152	Aau91152 Mycobacte	338	33	58.9	929	8	ADT04046	Adt04046 Human pro
266	33	58.9	281	8	ADK76931	Adk76931 Plant ful	339	33	58.9	932	4	ABG22206	Abg22206 Novel hum
267	33	58.9	281	8	AAU44866	Aau44866 Propionib	340	33	58.9	972	8	ADR10276	Adr10276 Human pro
268	33	58.9	287	4	ABM41385	Abm41385 Propionib	341	33	58.9	1044	4	ABG30271	Abg30271 Novel hum
269	33	58.9	287	6	ABM41385	Abm41385 Mycobacte	342	33	58.9	1047	4	ABG24636	Abg24636 Novel hum
270	33	58.9	287	7	ADB80169	Adb80169 Human met	343	33	58.9	1089	7	ADM29411	Adm29411 Human nov
271	33	58.9	290	4	ABU53127	Abu53127 Bacterial	344	33	58.9	1118	2	AAW82395	Aaw82395 Human UBP
272	33	58.9	290	4	ABU53127	Abu53127 Bacterial	345	33	58.9	1118	7	ADG10514	Adg10514 Human STA
273	33	58.9	332	3	ABU53447	Abu53447 Human col	346	33	58.9	1118	7	ADM29413	Adm29413 Human nov
274	33	58.9	332	3	ABU40148	Abu40148 Protein e	347	33	58.9	1118	7	ADG10514	Adg10514 Human STA
275	33	58.9	333	8	ADK14053	Adk14053 C. glutam	348	33	58.9	1118	9	ADH09574	Adh09574 Human nov
276	33	58.9	335	6	ADA8338	Ada8338 Rice prot	349	33	58.9	1128	4	ABG11843	Abg11843 Novel hum
277	33	58.9	342	8	ADK14055	Adk14055 C. glutam	350	33	58.9	1128	4	ABG11843	Abg11843 Novel hum
278	33	58.9	345	2	AAU31651	Aau31651 Corynebact	351	33	58.9	1130	4	ABG22207	Abg22207 Novel hum
279	33	58.9	345	2	AAU31651	Aau31651 Corynebact	352	33	58.9	1130	4	ABG22207	Abg22207 Novel hum
280	33	58.9	348	7	ABO77897	Abu77897 Pseudomon	353	33	58.9	1130	4	ABG22207	Abg22207 Novel hum
281	33	58.9	363	4	AG91433	Ag91433 C. glutami	354	32.5	58.0	173	7	ADD19312	Add19312 Human sec
282	33	58.9	366	6	ADB06642	Adb06642 Alloioacc	355	32.5	58.0	205	7	ADD19312	Add19312 Human sec
283	33	58.9	375	8	ADU35375	Adu35375 Haemophil	356	32.5	58.0	405	4	AAU38919	Aau38919 C. tracho
284	33	58.9	396	6	ABU30159	Abu30159 Protein e	357	32.5	58.0	405	7	ADU42710	Adu42710 Chlamydia
285	33	58.9	403	6	ABU44787	Abu44787 Protein e	358	32.5	58.0	566	3	AAU87134	Aau87134 Human sec
286	33	58.9	405	8	ADK72956	Adk72956 Plant ful	359	32.5	58.0	566	4	AAE06111	Aae06111 Human gen
287	33	58.9	423	7	ABO69948	Abu69948 Pseudomon	360	32.5	58.0	566	5	ABG33933	Abg33933 Human sec
288	33	58.9	441	7	ADG10512	Adg10512 Human STA	361	32.5	58.0	566	6	ADA57360	Ada57360 Human sec
289	33	58.9	444	7	ABO78759	Abu78759 Pseudomon	362	32.5	58.0	566	6	ADA41241	Ada41241 Human sec
290	33	58.9	448	5	ABB92481	Abb92481 Herbicida	363	32.5	58.0	566	6	ABR48002	Abu48002 Human sec
291	33	58.9	451	8	ADK14054	Adk14054 C. glutam	364	32.5	58.0	567	2	AAU75770	Aau75770 Human oxi
292	33	58.9	460	8	ADS26304	Ads26304 Bacterial	365	32.5	58.0	567	2	AAU75770	Aau75770 Human oxi
293	33	58.9	497	6	ADR06644	Adr06644 Alloioacc	366	32.5	58.0	567	2	AAU75770	Aau75770 Human oxi
294	33	58.9	507	4	ABG28559	Abu28559 Novel hum	367	32.5	58.0	567	3	AAU99364	Aau99364 Human PRO
295	33	58.9	513	6	ABU54545	Abu54545 Human NOV	368	32.5	58.0	567	3	AAU99364	Aau99364 Human PRO
296	33	58.9	517	4	AAU23711	Aau23711 Novel hum	369	32.5	58.0	567	4	AAU99364	Aau99364 Human PRO
297	33	58.9	530	2	AAU31649	Aau31649 Brevibact	370	32.5	58.0	567	4	AAU12409	Aau12409 Human PRO
298	33	58.9	530	2	AAU31649	Aau31649 Brevibact	371	32.5	58.0	567	5	ABB84904	Abb84904 Human PRO
299	33	58.9	530	2	AAU31649	Aau31649 Brevibact	372	32.5	58.0	567	5	ABB84904	Abb84904 Human PRO
300	33	58.9	530	4	AAU31649	Aau31649 Brevibact	373	32.5	58.0	567	5	ABB84904	Abb84904 Human PRO
301	33	58.9	547	8	ADN17962	Adn17962 Bacterial	374	32.5	58.0	567	6	ABO17853	Abu17853 Novel hum
302	33	58.9	553	8	ABM96964	Abm96964 M. xanthu	375	32.5	58.0	567	6	ABO17853	Abu17853 Novel hum
303	33	58.9	569	8	ADU07623	Adu07623 Amino aci	376	32.5	58.0	567	6	ABU56835	Abu56835 Human sec
304	33	58.9	592	8	ADK89308	Adk89308 Plant ful	377	32.5	58.0	567	6	ABU56835	Abu56835 Human sec
305	33	58.9	602	8	ADU09874	Adu09874 Plant ful	378	32.5	58.0	567	6	ABU66807	Abu66807 Human PRO
306	33	58.9	605	7	ABO77909	Abu77909 Pseudomon	379	32.5	58.0	567	6	ABU66807	Abu66807 Human PRO
307	33	58.9	608	7	ABO66646	Abu66646 Alloioacc	380	32.5	58.0	567	6	ADA40866	Ada40866 Human sec
308	33	58.9	608	7	ABO64692	Abu64692 Klebsiell	381	32.5	58.0	567	6	ABR47699	Abu47699 Human sec
309	33	58.9	622	8	ADN21005	Adn21005 Bacterial	382	32.5	58.0	567	6	ABU59888	Abu59888 Novel sec
310	33	58.9	634	7	ADG10510	Adg10510 Human STA	383	32.5	58.0	567	6	ABG74758	Abu74758 Human sec
311	33	58.9	634	7	ADG10510	Adg10510 Human STA	384	32.5	58.0	567	6	ABO25078	Abu25078 Human sec
312	33	58.9	650	6	ADB06648	Adb06648 Alloioacc	385	32.5	58.0	567	6	ABO25078	Abu25078 Human sec
313	33	58.9	666	8	ADK43108	Adk43108 Bacterial	386	32.5	58.0	567	6	ADA45995	Ada45995 Novel hum
314	33	58.9	672	8	ADK76084	Adk76084 Plant ful	387	32.5	58.0	567	6	ADA45995	Ada45995 Novel hum
315	33	58.9	673	8	ADL05485	Adl05485 M. catarr	388	32.5	58.0	567	6	ADA76426	Ada76426 Human PRO
316	33	58.9	682	6	ABU28397	Abu28397 Protein e	389	32.5	58.0	567	6	ADA61699	Ada61699 Homo sapi

390	32.5	58.0	567	6	ADB19484	Novel	hum	463	32.5	58.0	567	7	ADB36249	Human	PRO
391	32.5	58.0	567	6	ADB28025	Human	PRO	464	32.5	58.0	567	7	ADB46644	Novel	hum
392	32.5	58.0	567	6	ADA86504	Novel	hum	465	32.5	58.0	567	7	ADC17953	Human	PRO
393	32.5	58.0	567	6	ADB16068	Human	PRO	466	32.5	58.0	567	7	ADC50517	Novel	hum
394	32.5	58.0	567	6	ADA47854	Human	PRO	467	32.5	58.0	567	7	ADC72064	Novel	hum
395	32.5	58.0	567	6	ABO33606	Novel	hum	468	32.5	58.0	567	7	ADC60043	Novel	hum
396	32.5	58.0	567	6	ADA67649	Human	PRO	469	32.5	58.0	567	7	ADC53050	Novel	hum
397	32.5	58.0	567	6	ADB30656	Human	PRO	470	32.5	58.0	567	7	ADC57404	Novel	hum
398	32.5	58.0	567	6	ADA85952	Novel	hum	471	32.5	58.0	567	7	ADC60595	Novel	hum
399	32.5	58.0	567	6	ADA97164	Human	PRO	472	32.5	58.0	567	7	ADC51070	Novel	hum
400	32.5	58.0	567	6	ADA79468	Human	PRO	473	32.5	58.0	567	7	ADC65597	Human	PRO
401	32.5	58.0	567	6	ADA87607	Novel	hum	474	32.5	58.0	567	7	ADC54695	Novel	hum
402	32.5	58.0	567	6	ADB16809	Human	PRO	475	32.5	58.0	567	7	ADC53656	Novel	hum
403	32.5	58.0	567	6	ADA91901	Novel	hum	476	32.5	58.0	567	7	ADC59179	Novel	hum
404	32.5	58.0	567	6	ADB14964	Human	PRO	477	32.5	58.0	567	7	ADC56057	Novel	hum
405	32.5	58.0	567	6	ADB18925	Novel	hum	478	32.5	58.0	567	7	ADC58627	Novel	hum
406	32.5	58.0	567	6	ADA94140	Human	PRO	479	32.5	58.0	567	7	ADD03301	Novel	hum
407	32.5	58.0	567	6	ADB20036	Novel	hum	480	32.5	58.0	567	7	ADC90293	Novel	hum
408	32.5	58.0	567	6	ADB13348	Human	PRO	481	32.5	58.0	567	7	ADC69712	Human	PRO
409	32.5	58.0	567	6	ABO43386	Novel	hum	482	32.5	58.0	567	7	ADC48601	Human	PRO
410	32.5	58.0	567	6	ADA74602	Human	PRO	483	32.5	58.0	567	7	ADD10130	Human	PRO
411	32.5	58.0	567	6	ADB24835	Human	PRO	484	32.5	58.0	567	7	ADD04705	Novel	hum
412	32.5	58.0	567	6	ADA82359	Human	PRO	485	32.5	58.0	567	7	ADC80661	Novel	hum
413	32.5	58.0	567	6	ADA75322	Human	PRO	486	32.5	58.0	567	7	ADD11168	Human	PRO
414	32.5	58.0	567	6	ADA85400	Novel	hum	487	32.5	58.0	567	7	ADD10465	Human	sec
415	32.5	58.0	567	6	ADA84848	Novel	hum	488	32.5	58.0	567	7	ADC48049	Human	PRO
416	32.5	58.0	567	6	ADB30104	Human	PRO	489	32.5	58.0	567	7	ADC80109	Novel	hum
417	32.5	58.0	567	6	ADA80632	Human	PRO	490	32.5	58.0	567	7	ADD11425	Human	sec
418	32.5	58.0	567	6	ADA75874	Human	PRO	491	32.5	58.0	567	7	ADD09578	Human	PRO
419	32.5	58.0	567	6	ADA47099	Human	PRO	492	32.5	58.0	567	7	ADD41291	Novel	hum
420	32.5	58.0	567	6	ADB25395	Human	PRO	493	32.5	58.0	567	7	ADD52430	Human	PRO
421	32.5	58.0	567	6	ADA93571	Human	PRO	494	32.5	58.0	567	7	ADD15372	Novel	hum
422	32.5	58.0	567	6	ADB26921	Human	PRO	495	32.5	58.0	567	7	ADD70599	Human	sec
423	32.5	58.0	567	6	ADB31208	Human	PRO	496	32.5	58.0	567	7	ADD39676	Human	sec
424	32.5	58.0	567	6	ADA61136	Homo sapi		497	32.5	58.0	567	7	ADD53170	Human	PRO
425	32.5	58.0	567	6	ADB24283	Human	PRO	498	32.5	58.0	567	7	ADD53722	Novel	hum
426	32.5	58.0	567	6	ADA96612	Human	PRO	499	32.5	58.0	567	7	ADD70122	Human	sec
427	32.5	58.0	567	6	ADA81184	Human	PRO	500	32.5	58.0	567	7	ADD37218	Human	sec
428	32.5	58.0	567	6	ADA96060	Human	PRO	501	32.5	58.0	567	7	ADD38243	Human	sec
429	32.5	58.0	567	6	ADB26369	Human	PRO	502	32.5	58.0	567	7	ADD39199	Human	sec
430	32.5	58.0	567	6	ADB21854	Novel	hum	503	32.5	58.0	567	7	ADD51878	Human	PRO
431	32.5	58.0	567	7	ADA77633	Human	PRO	504	32.5	58.0	567	7	ADD02677	Human	PRO
432	32.5	58.0	567	7	ADB18373	Human	PRO	505	32.5	58.0	567	7	ADD02111	Human	PRO
433	32.5	58.0	567	7	ADA87056	Novel	hum	506	32.5	58.0	567	7	ADD54293	Novel	hum
434	32.5	58.0	567	7	ABO44459	Human	sec	507	32.5	58.0	567	7	ADD38722	Human	sec
435	32.5	58.0	567	7	ADA88159	Novel	hum	508	32.5	58.0	567	7	ADD40153	Human	sec
436	32.5	58.0	567	7	ADA46547	Novel	hum	509	32.5	58.0	567	7	ADD50374	Human	sec
437	32.5	58.0	567	7	ADB28577	Human	PRO	510	32.5	58.0	567	7	ADD92610	Human	PRO
438	32.5	58.0	567	7	ADB29129	Human	PRO	511	32.5	58.0	567	7	ADD91506	Human	PRO
439	32.5	58.0	567	7	ADA77081	Human	PRO	512	32.5	58.0	567	7	ADB04120	Human	PRO
440	32.5	58.0	567	7	ABO33483	Novel	hum	513	32.5	58.0	567	7	ADB19986	Human	sec
441	32.5	58.0	567	7	ADA88711	Novel	hum	514	32.5	58.0	567	7	ADB32417	Novel	hum
442	32.5	58.0	567	7	ADA97716	Human	PRO	515	32.5	58.0	567	7	ADB22349	Human	PRO
443	32.5	58.0	567	7	ADB27473	Human	PRO	516	32.5	58.0	567	7	ADD79573	Human	PRO
444	32.5	58.0	567	7	ADB22406	Novel	hum	517	32.5	58.0	567	7	ADB42109	Human	PRO
445	32.5	58.0	567	7	ADA67097	Human	PRO	518	32.5	58.0	567	7	ADB17926	Human	PRO
446	32.5	58.0	567	7	ADB22958	Human	PRO	519	32.5	58.0	567	7	ADD92058	Human	PRO
447	32.5	58.0	567	7	ADB23731	Human	PRO	520	32.5	58.0	567	7	ADB33521	Novel	hum
448	32.5	58.0	567	7	ADA92453	Novel	hum	521	32.5	58.0	567	7	ADB34073	Novel	hum
449	32.5	58.0	567	7	ADB15516	Human	PRO	522	32.5	58.0	567	7	ADD80125	Human	PRO
450	32.5	58.0	567	7	ADB38768	Novel	hum	523	32.5	58.0	567	7	ADB49897	Human	sec
451	32.5	58.0	567	7	ADB38216	Novel	hum	524	32.5	58.0	567	7	ADB193162	Human	PRO
452	32.5	58.0	567	7	ADB66688	Novel	hum	525	32.5	58.0	567	7	ADB19582	Human	PRO
453	32.5	58.0	567	7	ADB89768	Human	PRO	526	32.5	58.0	567	7	ADB21455	Human	sec
454	32.5	58.0	567	7	ADB90500	Human	PRO	527	32.5	58.0	567	7	ADB19030	Human	PRO
455	32.5	58.0	567	7	ADB39601	Novel	hum	528	32.5	58.0	567	7	ADB43226	Human	PRO
456	32.5	58.0	567	7	ADB47224	Novel	hum	529	32.5	58.0	567	7	ADB96015	Human	PRO
457	32.5	58.0	567	7	ADB86831	Human	PRO	530	32.5	58.0	567	7	ADB22901	Human	PRO
458	32.5	58.0	567	7	ADB77436	Novel	hum	531	32.5	58.0	567	7	ADD79019	Human	PRO
459	32.5	58.0	567	7	ADB34593	Human	PRO	532	32.5	58.0	567	7	ADB32969	Novel	hum
460	32.5	58.0	567	7	ADB35697	Human	PRO	533	32.5	58.0	567	7	ADB42661	Human	PRO
461	32.5	58.0	567	7	ADB34041	Human	PRO	534	32.5	58.0	567	7	ADD80677	Human	PRO
462	32.5	58.0	567	7	ADB35145	Human	PRO	535	32.5	58.0	567	7	ADD89705	Human	PRO

536	32.5	58.0	567	7	Ade40989	Human	PRO	Ad62121	Novel	hum	609	32.5	58.0	567	8	ADG62121	Novel	hum	Ad62121	Novel	hum
537	32.5	58.0	567	7	Ad804788	Human	PRO	AdH02972	Human	sec	610	32.5	58.0	567	8	ADH02972	Human	sec	AdH02972	Human	sec
538	32.5	58.0	567	7	Ad929117	Human	PRO	AdG82322	Novel	hum	611	32.5	58.0	567	8	ADG82322	Novel	hum	AdG82322	Novel	hum
539	32.5	58.0	567	7	Adf29880	Human	sec	AdG57561	Novel	hum	612	32.5	58.0	567	8	ADG57561	Novel	hum	AdG57561	Novel	hum
540	32.5	58.0	567	7	Adf94077	Immune di		AdG57009	Novel	hum	613	32.5	58.0	567	8	ADG57009	Novel	hum	AdG57009	Novel	hum
541	32.5	58.0	567	7	Adf55773	Human	sec	AdG55905	Novel	hum	614	32.5	58.0	567	8	ADG55905	Novel	hum	AdG55905	Novel	hum
542	32.5	58.0	567	7	Adg21626	Novel	hum	AdG58665	Novel	hum	615	32.5	58.0	567	8	ADG58665	Novel	hum	AdG58665	Novel	hum
543	32.5	58.0	567	7	Adg23267	Novel	hum	AdG71031	Novel	hum	616	32.5	58.0	567	8	ADG71031	Novel	hum	AdG71031	Novel	hum
544	32.5	58.0	567	7	Adp97602	Human	PRO	AdH03449	Human	sec	617	32.5	58.0	567	8	ADH03449	Human	sec	AdH03449	Human	sec
545	32.5	58.0	567	7	Adg80666	Human	PRO	AdG58113	Novel	hum	618	32.5	58.0	567	8	ADG58113	Novel	hum	AdG58113	Novel	hum
546	32.5	58.0	567	7	Adg80114	Human	PRO	AdG53697	Novel	hum	619	32.5	58.0	567	8	ADG53697	Novel	hum	AdG53697	Novel	hum
547	32.5	58.0	567	7	Adh55406	Novel	hum	AdG71583	Novel	hum	620	32.5	58.0	567	8	ADG71583	Novel	hum	AdG71583	Novel	hum
548	32.5	58.0	567	7	Adh55958	Novel	hum	AdG81770	Human	PRO	621	32.5	58.0	567	8	ADG81770	Human	PRO	AdG81770	Human	PRO
549	32.5	58.0	567	7	Adh99277	Human	sec	AdH30732	Human	PRO	622	32.5	58.0	567	8	ADH30732	Human	PRO	AdH30732	Human	PRO
550	32.5	58.0	567	7	Adl65126	Novel	hum	AdH12099	Novel	hum	623	32.5	58.0	567	8	ADH12099	Novel	hum	AdH12099	Novel	hum
551	32.5	58.0	567	7	Adl63625	Novel	hum	AdG52521	Novel	hum	624	32.5	58.0	567	8	ADG52521	Novel	hum	AdG52521	Novel	hum
552	32.5	58.0	567	7	Adh82039	Novel	hum	AdG54249	Novel	hum	625	32.5	58.0	567	8	ADG54249	Novel	hum	AdG54249	Novel	hum
553	32.5	58.0	567	7	Adh81487	Novel	hum	AdG581218	Human	PRO	626	32.5	58.0	567	8	ADG581218	Human	PRO	AdG581218	Human	PRO
554	32.5	58.0	567	7	Adn16055	Novel	hum	AdG56457	Novel	hum	627	32.5	58.0	567	8	ADG56457	Novel	hum	AdG56457	Novel	hum
555	32.5	58.0	567	7	Adn16684	Novel	hum	AdH12723	Novel	hum	628	32.5	58.0	567	8	ADH12723	Novel	hum	AdH12723	Novel	hum
556	32.5	58.0	567	7	Adn15503	Novel	hum	AdG61569	Novel	hum	629	32.5	58.0	567	8	ADG61569	Novel	hum	AdG61569	Novel	hum
557	32.5	58.0	567	7	Adn14951	Novel	hum	AdH28656	Human	PRO	630	32.5	58.0	567	8	ADH28656	Human	PRO	AdH28656	Human	PRO
558	32.5	58.0	567	7	Adi64177	Novel	hum	AdG54801	Novel	hum	631	32.5	58.0	567	8	ADG54801	Novel	hum	AdG54801	Novel	hum
559	32.5	58.0	567	7	Adc81213	Novel	hum	AdG59841	Novel	hum	632	32.5	58.0	567	8	ADG59841	Novel	hum	AdG59841	Novel	hum
560	32.5	58.0	567	7	Adc81213	Novel	hum	AdH43609	Human	PRO	633	32.5	58.0	567	8	ADH43609	Human	PRO	AdH43609	Human	PRO
561	32.5	58.0	567	8	AdD76661	Human	PRO	AdH04403	Human	sec	634	32.5	58.0	567	8	ADH04403	Human	sec	AdH04403	Human	sec
562	32.5	58.0	567	8	AdD88025	Human	PRO	AdH61404	Human	sec	635	32.5	58.0	567	8	ADH61404	Human	sec	AdH61404	Human	sec
563	32.5	58.0	567	8	AdD86429	Human	PRO	AdH15479	Novel	hum	636	32.5	58.0	567	8	ADH15479	Novel	hum	AdH15479	Novel	hum
564	32.5	58.0	567	8	AdE75877	Human	sec	AdG10008	Novel	hum	637	32.5	58.0	567	8	ADG10008	Novel	hum	AdG10008	Novel	hum
565	32.5	58.0	567	8	AdE41426	Human	sec	AdG09336	Novel	hum	638	32.5	58.0	567	8	ADG09336	Novel	hum	AdG09336	Novel	hum
566	32.5	58.0	567	8	AdE23453	Human	PRO	AdI14811	Novel	hum	639	32.5	58.0	567	8	ADI14811	Novel	hum	AdI14811	Novel	hum
567	32.5	58.0	567	8	AdE24005	Human	PRO	AdI18406	Novel	hum	640	32.5	58.0	567	8	ADI18406	Novel	hum	AdI18406	Novel	hum
568	32.5	58.0	567	8	AdE24648	Human	PRO	AdK52108	Human	ato	641	32.5	58.0	567	8	ADK52108	Human	ato	AdK52108	Human	ato
569	32.5	58.0	567	8	AdE87473	Human	PRO	AdJ36887	Novel	hum	642	32.5	58.0	567	8	ADJ36887	Novel	hum	AdJ36887	Novel	hum
570	32.5	58.0	567	8	AdE89339	Human	PRO	AdJ77582	Human	PRO	643	32.5	58.0	567	8	ADJ77582	Human	PRO	AdJ77582	Human	PRO
571	32.5	58.0	567	8	AdE18478	Human	PRO	AdK82954	Human	PRO	644	32.5	58.0	567	8	ADK82954	Human	PRO	AdK82954	Human	PRO
572	32.5	58.0	567	8	AdE88787	Human	PRO	AdM27840	Human	PRO	645	32.5	58.0	567	8	ADM27840	Human	PRO	AdM27840	Human	PRO
573	32.5	58.0	567	8	AdE94807	Human	PRO	AdJ65704	Human	PRO	646	32.5	58.0	567	8	ADJ65704	Human	PRO	AdJ65704	Human	PRO
574	32.5	58.0	567	8	AdE91218	Human	PRO	AdM7807	Human	PRO	647	32.5	58.0	567	8	ADM7807	Human	PRO	AdM7807	Human	PRO
575	32.5	58.0	567	8	AdE91218	Human	PRO	AdM42564	Human	PRO	648	32.5	58.0	567	8	ADM42564	Human	PRO	AdM42564	Human	PRO
576	32.5	58.0	567	8	Adf25768	Human	sec	AdI91823	Human	PRO	649	32.5	58.0	567	8	ADI91823	Human	PRO	AdI91823	Human	PRO
577	32.5	58.0	567	8	AdE95359	Human	PRO	AdI94603	Human	sec	650	32.5	58.0	567	8	ADI94603	Human	sec	AdI94603	Human	sec
578	32.5	58.0	567	8	AdE93469	Human	PRO	AdM28426	Human	PRO	651	32.5	58.0	567	8	ADM28426	Human	PRO	AdM28426	Human	PRO
579	32.5	58.0	567	8	Adf24667	Human	sec	AdI95908	Human	PRO	652	32.5	58.0	567	8	ADI95908	Human	PRO	AdI95908	Human	PRO
580	32.5	58.0	567	8	Adf29403	Human	PRO	AdI96460	Novel	hum	653	32.5	58.0	567	8	ADI96460	Novel	hum	AdI96460	Novel	hum
581	32.5	58.0	567	8	Adf35050	Human	PRO	AdP55860	Human	PRO	654	32.5	58.0	567	8	ADP55860	Human	PRO	AdP55860	Human	PRO
582	32.5	58.0	567	8	AdE96934	Human	sec	AdS32412	Novel	hum	655	32.5	58.0	567	8	ADS32412	Novel	hum	AdS32412	Novel	hum
583	32.5	58.0	567	8	AdE92365	Novel	hum	AdT03396	Human	PRO	656	32.5	58.0	567	8	ADT03396	Human	PRO	AdT03396	Human	PRO
584	32.5	58.0	567	8	AdE90666	Human	PRO	AdX7621	Rheumatoi		657	32.5	58.0	567	9	ADX7621	Rheumatoi		AdX7621	Rheumatoi	
585	32.5	58.0	567	8	AdG91813	Novel	hum	AdB14193	Cancer ce		658	32.5	58.0	567	9	ADB14193	Cancer ce		AdB14193	Cancer ce	
586	32.5	58.0	567	8	AdG02392	Human	PRO	AdY22368	TPO recep		659	32.5	58.0	567	9	ADY22368	TPO recep		AdY22368	TPO recep	
587	32.5	58.0	567	8	AdG22178	Novel	hum	AdA19908	Residues		660	32.5	58.0	567	9	ADA19908	Residues		AdA19908	Residues	
588	32.5	58.0	567	8	AdG20248	Human	PRO	AdA97176	Peptide d		661	32.5	58.0	567	9	ADA97176	Peptide d		AdA97176	Peptide d	
589	32.5	58.0	567	8	Adf98154	Human	PRO	AdA17396	CAP37 mim		662	32.5	58.0	567	9	ADA17396	CAP37 mim		AdA17396	CAP37 mim	
590	32.5	58.0	567	8	AdG24371	Novel	hum	AdA909889	Cationic		663	32.5	58.0	567	9	ADA909889	Cationic		AdA909889	Cationic	
591	32.5	58.0	567	8	AdG98725	Human	PRO	AdB73357	Exemplary		664	32.5	58.0	567	9	ADB73357	Exemplary		AdB73357	Exemplary	
592	32.5	58.0	567	8	AdG03556	Human	PRO	AdG42093	Human	CAP	665	32.5	58.0	567	9	ADG42093	Human	CAP	AdG42093	Human	CAP
593	32.5	58.0	567	8	Adf99277	Human	PRO	AdJ73510	Exemplary		666	32.5	58.0	567	9	ADJ73510	Exemplary		AdJ73510	Exemplary	
594	32.5	58.0	567	8	AdG16862	Human	PRO	AdR41925	Internal		667	32.5	58.0	567	9	ADR41925	Internal		AdR41925	Internal	
595	32.5	58.0	567	8	AdG05321	Human	PRO	AdR67974	Fragment		668	32.5	58.0	567	9	ADR67974	Fragment		AdR67974	Fragment	
596	32.5	58.0	567	8	AdG19588	Human	PRO	AdR33764	Cowpea ge		669	32.5	58.0	567	9	ADR33764	Cowpea ge		AdR33764	Cowpea ge	
597	32.5	58.0	567	8	AdG13425	Human	PRO	AdP00712	Human	ORF	670	32.5	58.0	567	9	ADP00712	Human	ORF	AdP00712	Human	ORF
598	32.5	58.0	567	8	AdG08482	Novel	hum	AdP01989	Human	ORF	671	32.5	58.0	567	9	ADP01989	Human	ORF	AdP01989	Human	ORF
599	32.5	58.0	567	8	AdG15652	Human	PRO	AdB28733	Arabidops		672	32.5	58.0	567	9	ADB28733	Arabidops		AdB28733	Arabidops	
600	32.5	58.0	567	8	AdG15652	Human	PRO	AdB080350	Pseudomon		673	32.5	58.0	567	9	ADB080350	Pseudomon		AdB080350	Pseudomon	
601	32.5	58.0	567	8	AdG06235	Human	PRO	AdB76670	Human	col	674	32.5	58.0	567	9	ADB76670	Human	col	AdB76670	Human	col
602	32.5	58.0	567	8	AdG23819	Novel	hum	AdX10411	Human	pol	675	32.5	58.0	567	9	ADX10411	Human	pol	AdX10411	Human	pol
603	32.5	58.0	567	8	AdG04108	Human	PRO	AdX89010	Plant fu		676	32.5	58.0	567	9	ADX89010	Plant fu		AdX89010	Plant fu	
604	32.5	58.0	567	8	AdG04108	Human	PRO				677</										

682	32	57.1	98	8	ADK69828	Adk69828	TIRAP pro	755	32	57.1	249	8	ADK93328	Adk93328	Plant ful
683	32	57.1	103	3	AAB42523	Aab42523	Human ORF	756	32	57.1	250	3	AAG18126	Aag18126	Arabidops
684	32	57.1	118	4	AUA47072	Aua47072	Propionib	757	32	57.1	251	2	AAR10668	Aar10668	Cationic
685	32	57.1	118	6	ABM43591	Abm43591	Propionib	758	32	57.1	251	2	AAR84666	Aar84666	Human hep
686	32	57.1	124	3	AAG30311	Aag30311	Arabidops	759	32	57.1	251	2	AAR88365	Aar88365	Human pre
687	32	57.1	124	3	AAG50533	Aag50533	Arabidops	760	32	57.1	251	2	AAW88121	Aaw88121	Complete
688	32	57.1	129	3	AAG30101	Aag30101	Arabidops	761	32	57.1	251	3	AAW71878	Aay71878	Human pre
689	32	57.1	129	3	AAG50532	Aag50532	Arabidops	762	32	57.1	251	4	AAW71894	Aay71894	Human pre
690	32	57.1	136	3	AAG35509	Aag35509	Arabidops	763	32	57.1	251	9	ADK69318	Adk69318	Human hep
691	32	57.1	143	3	AAV95003	Aay95003	Human sec	764	32	57.1	251	3	AAG53308	Aag53308	Arabidops
692	32	57.1	143	4	AAB95555	Aab95555	Human pro	765	32	57.1	261	3	AAG06132	Aag06132	Arabidops
693	32	57.1	143	4	ABG02114	Abg02114	Novel hum	766	32	57.1	262	8	ADN25465	Adn25465	Bacterial
694	32	57.1	143	5	ABB98073	Abb98073	Human ary	767	32	57.1	262	8	ADY08479	Ady08479	Plant ful
695	32	57.1	143	6	ABU06886	Abu06886	Maize SSI	768	32	57.1	265	6	ABU36258	Abu36258	Pseudomon
696	32	57.1	146	6	ABU06685	Abu06685	Maize SSI	769	32	57.1	265	6	ABU38406	Abu38406	Protein e
697	32	57.1	147	3	AAG48729	Aag48729	Arabidops	770	32	57.1	266	8	ADY23109	Ady23109	Plant ful
698	32	57.1	147	3	AAG05246	Aag05246	Arabidops	771	32	57.1	272	7	ABO81531	Abo81531	Pseudomon
699	32	57.1	154	3	AAG07540	Aag07540	Arabidops	772	32	57.1	274	2	AAR84665	Aar84665	Human hep
700	32	57.1	154	5	ABG60122	Abg60122	Human DIT	773	32	57.1	277	6	ABM64806	Abm64806	Propionib
701	32	57.1	161	8	ADK90415	Adk90415	Plant ful	774	32	57.1	288	6	ABR42539	Abra42539	Clorobioc
702	32	57.1	165	8	ADT57682	Adt57682	Plant pol	775	32	57.1	289	8	ADY08001	Ady08001	Plant ful
703	32	57.1	167	7	ABO76469	Abo76469	Pseudomon	776	32	57.1	290	8	ADY73222	Ady73222	Plant ful
704	32	57.1	171	3	AAG35508	Aag35508	Arabidops	777	32	57.1	301	3	AAG13112	Aag13112	Arabidops
705	32	57.1	171	6	ABU43429	Abu43429	Protein e	778	32	57.1	309	3	AAG06131	Aag06131	Arabidops
706	32	57.1	172	6	ABU39620	Abu39620	Protein e	779	32	57.1	314	4	AAB70085	Aab70085	Human sec
707	32	57.1	182	3	AAG48728	Aag48728	Arabidops	780	32	57.1	314	8	ADY12461	Ady12461	Plant ful
708	32	57.1	182	3	AAG05245	Aag05245	Arabidops	781	32	57.1	319	3	AAG39923	Aag39923	Arabidops
709	32	57.1	182	5	ABB93582	Abb93582	Herbicida	782	32	57.1	321	3	AAG53307	Aag53307	Arabidops
710	32	57.1	185	4	AAU32997	Aau32997	Novel hum	783	32	57.1	324	3	AAG07122	Aag07122	Arabidops
711	32	57.1	191	3	AAG48727	Aag48727	Arabidops	784	32	57.1	326	8	ADK95574	Adk95574	Plant ful
712	32	57.1	191	3	AAG35507	Aag35507	Arabidops	785	32	57.1	328	8	ADY74803	Ady74803	Plant ful
713	32	57.1	192	3	AAG05244	Aag05244	Arabidops	786	32	57.1	329	3	AAG06130	Aag06130	Arabidops
714	32	57.1	197	1	AAP91931	Aap91931	Human hep	787	32	57.1	329	4	AAW78574	Aam78574	Human pro
715	32	57.1	198	3	AAB56131	Aab56131	Human sec	788	32	57.1	329	8	ADK97122	Adk97122	Plant ful
716	32	57.1	199	7	ADF17970	Adf17970	Human cat	789	32	57.1	330	8	ADK95575	Adk95575	Plant ful
717	32	57.1	199	7	ADG42089	Adg42089	Human cor	790	32	57.1	332	5	ABB08371	Abb08371	Y-FVM*7 a
718	32	57.1	200	9	ADV85561	Adv85561	Mutant ad	791	32	57.1	335	2	AAW15420	Aaw15420	Rat Sprm-
719	32	57.1	207	7	ADC30968	Adc30968	Human nov	792	32	57.1	340	8	ADK24219	Adk24219	Bacterial
720	32	57.1	212	4	ABG23125	Abg23125	Novel hum	793	32	57.1	347	8	ADJ35032	Adj35032	Xylanase
721	32	57.1	216	8	ADX75990	Adx75990	Plant ful	794	32	57.1	349	8	ADS24608	Ads24608	Bacterial
722	32	57.1	217	6	AAB06500	Aab06500	Maize SSI	795	32	57.1	350	4	AAU61038	Aau61038	Propionib
723	32	57.1	221	2	AAR84663	Aar84663	Human hep	796	32	57.1	350	6	ABM57557	Abm57557	Propionib
724	32	57.1	221	2	AAW04875	Aaw04875	Heparin b	797	32	57.1	352	4	ABG08205	Abg08205	Novel hum
725	32	57.1	221	8	ADT57175	Adt57175	Plant pol	798	32	57.1	352	6	ABM68224	Abm68224	Photorhab
726	32	57.1	222	2	AAK41935	Aak41935	Recombina	799	32	57.1	352	3	AAG53306	Aag53306	Arabidops
727	32	57.1	222	2	AAW73210	Aaw73210	CAP37 pro	800	32	57.1	360	3	AAG07121	Aag07121	Arabidops
728	32	57.1	222	7	ADG42088	Adg42088	Human pol	801	32	57.1	360	7	ADE25082	Ade25082	Plant gro
729	32	57.1	225	2	AAW88362	Aaw88362	Human mat	802	32	57.1	361	7	ADJ70843	Adj70843	Human hea
730	32	57.1	225	2	AAV21551	Aay21551	Human hep	803	32	57.1	362	8	ADK87805	Adk87805	Plant ful
731	32	57.1	225	2	AAW88118	Aaw88118	Mature hu	804	32	57.1	363	8	ADT56787	Adt56787	Plant pol
732	32	57.1	225	3	AAV71881	Aay71881	Human hep	805	32	57.1	364	3	AAV44258	Aay44258	Reversibl
733	32	57.1	225	3	AAV71882	Aay71882	Human hep	806	32	57.1	366	2	AAV37453	Aay37453	Amino aci
734	32	57.1	225	3	AAV71883	Aay71883	Human hep	807	32	57.1	366	2	AAV35499	Aay35499	Chlamydia
735	32	57.1	225	3	AAV71876	Aay71876	Human mat	808	32	57.1	366	6	ABU27052	Abu27052	Protein e
736	32	57.1	225	4	AAV71891	Aay71891	Human mat	809	32	57.1	366	6	ABU27139	Abu27139	Protein e
737	32	57.1	225	7	ADE11587	Ade11587	Human mat	810	32	57.1	368	8	ADY24592	Ady24592	Plant ful
738	32	57.1	225	7	ADE11551	Ade11551	Human mat	811	32	57.1	368	8	ADY75138	Ady75138	Plant ful
739	32	57.1	225	8	ADK42019	Adk42019	Human hep	812	32	57.1	369	7	ADC68413	Adc68413	Liolum pe
740	32	57.1	225	9	ADV21099	Adv21099	Human azu	813	32	57.1	369	8	ADJ34800	Adj34800	Xylanase
741	32	57.1	225	9	ADK69310	Adk69310	Human hep	814	32	57.1	369	9	AEB03116	Aeb03116	Tannin bi
742	32	57.1	225	9	ADY82139	Ady82139	Human hep	815	32	57.1	371	5	ABB93226	Abb93226	Herbicida
743	32	57.1	227	5	AAU82726	Aau82726	Amino aci	816	32	57.1	372	8	ADN26117	Adn26117	Bacterial
744	32	57.1	228	5	ABG70077	Abg70077	Human pre	817	32	57.1	372	8	ADK67633	Adk67633	Plant ful
745	32	57.1	228	7	ADE11545	Ade11545	Human mat	818	32	57.1	372	8	ADK68611	Adk68611	Plant ful
746	32	57.1	230	3	AAG50531	Aag50531	Arabidops	819	32	57.1	372	8	ADY05197	Ady05197	Plant ful
747	32	57.1	230	3	AAG13009	Aag13009	Arabidops	820	32	57.1	372	8	ADY09421	Ady09421	Plant ful
748	32	57.1	232	2	AAW88364	Aaw88364	Human pro	821	32	57.1	372	9	ABM96610	Abm96610	M. xanthu
749	32	57.1	232	2	AAV21550	Aay21550	Human hep	822	32	57.1	372	8	ABY92248	Abby92248	Herbicida
750	32	57.1	232	3	AAW88120	Aaw88120	Human hep	823	32	57.1	373	5	ADY25279	Ady25279	Plant ful
751	32	57.1	232	3	AAV71877	Aay71877	Human pre	824	32	57.1	374	5	ADY25229	Ady25229	Plant ful
752	32	57.1	236	7	ABO83170	Abo83170	Pseudomon	825	32	57.1	376	3	AAG07120	Aag07120	Arabidops
753	32	57.1	244	4	AAV71893	Aay71893	Human pre	826	32	57.1	376	8	ADT49303	Adt49303	Pectate 1
754	32	57.1	245	8	ADN20995	Adn20995	Bacterial	827	32	57.1	378	5	ABB93127	Abb93127	Herbicida

974 32 57.1 915 6 ABO32762 Human sec
 975 32 57.1 915 6 ADA67467 Human PRO
 976 32 57.1 915 6 ADB30474 Human PRO
 977 32 57.1 915 6 ADA85770 Novel hum
 978 32 57.1 915 6 ADA96982 Human PRO
 979 32 57.1 915 6 ADA96982 Human PRO
 980 32 57.1 915 6 ADA87425 Novel hum
 981 32 57.1 915 6 ADB16627 Human PRO
 982 32 57.1 915 6 ABO34822 Human PRO
 983 32 57.1 915 6 ADA16070 Human sec
 984 32 57.1 915 6 ADA91719 Novel hum
 985 32 57.1 915 6 ADB14782 Human PRO
 986 32 57.1 915 6 ADB18743 Novel hum
 987 32 57.1 915 6 ADA93958 Human PRO
 988 32 57.1 915 6 ADB19854 Novel hum
 989 32 57.1 915 6 ADB13166 Novel hum
 990 32 57.1 915 6 ABO43295 Novel hum
 991 32 57.1 915 6 ADA74420 Human PRO
 992 32 57.1 915 6 ADA42215 Human sec
 993 32 57.1 915 6 ADB24653 Human PRO
 994 32 57.1 915 6 ADA82177 Human PRO
 995 32 57.1 915 6 ADA75140 Human PRO
 996 32 57.1 915 6 ADA85218 Novel hum
 997 32 57.1 915 6 ADA84666 Novel hum
 998 32 57.1 915 6 ABO17500 Human PRO
 999 32 57.1 915 6 ADB29922 Human PRO
 1000 32 57.1 915 6 ADA80450 Human PRO

ALIGNMENTS

RESULT 1
 AAY22355
 ID AAY22355 standard; peptide; 10 AA.
 XX
 AC AAY22355;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE TPO receptor binding peptide sequence, SEQ ID NO. 6.
 XX
 XX TPO: thrombopoietin receptor; thrombopoietin agonist; thrombocytopoenaia;
 KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
 KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
 XX
 OS Synthetic.
 XX
 XX US5932546-A.
 PN
 XX
 PD 03-AUG-1999.
 XX
 XX 04-OCT-1996; 96US-00726464.
 PF
 XX 04-OCT-1996; 96US-00726464.
 PR
 XX (GLAX) GLAXO WELLCOME INC.
 PA
 XX Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
 PI Cwirla SE, Johnson SS;
 PI
 XX WPI; 1999-457122/38.
 DR
 XX New low molecular weight thrombopoietin agonists, particularly peptides,
 PT for treatment of hematological disease and thrombocytopoenaia.
 PT
 XX Disclosure; Col 13-14; 36pp; English.
 PS
 XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
 CC peptide of the invention. The peptide has: (i) a molecular weight below
 CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
 CC nM. The peptides are used to treat conditions requiring a thrombopoietin
 CC agonist, particularly haematological disorders or thrombocytopoenaia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. Also when labelled they may be used for diagnosis
 CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
 CC in vitro expansion of megakaryocytes and committed progenitor cells, and
 CC for the development/identification of other TR agonists. The compounds
 CC accelerate thrombocyte regeneration
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 56; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00088;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GRVRDQVAGW 10
 Db 1 GRVRDQVAGW 10
 RESULT 2
 AAB16979
 ID AAB16979 standard; peptide; 10 AA.
 XX
 AC AAB16979;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE TPO-mimetic peptide sequence SEQ ID NO:35.
 XX
 XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 XX WO2000024782-A2.
 PN
 XX 04-MAY-2000.
 PD
 XX 25-OCT-1999; 99WO-US025044.
 PF
 XX 23-OCT-1998; 98US-0105371P.
 PR
 XX 22-OCT-1999; 99US-00428082.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 XX Feige U, Liu C, Cheetham J, Boone TC;
 PI
 XX WPI; 2000-350702/30.
 DR
 XX Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.
 PT
 XX Claim 19; Page 207; 608pp; English.
 PS
 XX The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to
 CC AAB19003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00088;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 |||||
 Db 1 GRVRDQVAGW 10

RESULT 3
 ABB72865
 ID ABB72865 standard; peptide; 10 AA.

XX
 AC ABB72865;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE TPO mimetic peptide SEQ ID NO:35.

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
 KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX
 XX Homo sapiens.
 OS Synthetic.
 OS
 XX WO200183525-A2.
 PN
 XX 08-NOV-2001.
 PD
 XX 02-MAY-2001; 2001WO-US014310.
 XX
 XX 03-MAY-2000; 2000US-00563286.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 PI
 XX WPI; 2002-130313/17.
 DR
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 PT
 XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.00088;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 |||||
 Db 1 GRVRDQVAGW 10

RESULT 4
 ADJ73016
 ID ADJ73016 standard; peptide; 10 AA.

XX
 AC ADJ73016;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE TPO mimetic peptide sequence SeqID 470.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW TPO.

XX
 OS Synthetic.
 OS
 XX WO2003084477-A2.
 PN
 XX 16-OCT-2003.
 PD
 XX 24-MAR-2003; 2003WO-US009139.
 XX
 XX 29-MAR-2002; 2002US-0368791P.
 XX
 XX (CENZ) CENTOCOR INC.
 XX
 XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
 PI
 XX WPI; 2003-804237/75.
 DR
 XX New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 470; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 56; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | | | | | | |
Db 1 GRVRDQVAGW 10

RESULT 5
ID ADJ52651
ADJ52651 standard; peptide; 10 AA.

XX AC ADJ52651;
XX 06-MAY-2004 (first entry)
XX CHI deleted mimetibody-related peptide SeqID470.

XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nontropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.
OS Synthetic.

XX WO2004002417-A2.
XX 08-JAN-2004.
XX 27-JUN-2003; 2003WO-US020347.
XX 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghraieb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
DR WPI; 2004-082870/08.

XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 470; 129pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nontropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 56; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
| | | | | | | | | |
Db 1 GRVRDQVAGW 10

RESULT 6
ID ADJ51612
ADJ51612 standard; peptide; 10 AA.

XX AC ADJ51612;
XX 06-MAY-2004 (first entry)
XX CHI deleted mimetibody-related peptide SeqID470.

XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW gastritis disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX Unidentified.
OS Synthetic.

XX WO2004002424-A2.
XX 08-JAN-2004.
XX 30-JUN-2003; 2003WO-US020495.
XX 28-JUN-2002; 2002US-0392431P.
XX 19-SEP-2002; 2002US-0412144P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghraieb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
DR WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic, and
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.

XX Claim 14; SEQ ID NO 470; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.
 XX
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.00088;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10

Db 1 GRVRDQVAGW 10

RESULT 7

ADN20123
 ID ADN20123 standard; protein; 780 AA.

XX

AC ADN20123;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #2776.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

XX

OS Bacteria.

XX

PN US2003233675-A1.

XX

PD 18-DEC-2003.

XX

PF 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX

DR WPI; 2004-061375/06.

XX

PT New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

XX

PS Claim 1; SEQ ID NO 2776; 122pp; English.

XX

CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 780 AA;

Query Match 73.2%; Score 41; DB 8; Length 780;

Best Local Similarity 77.8%; Pred. No. 58;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10

Db 357 RLADQVAGW 365

RESULT 8

ADA36768

ID ADA36768 standard; protein; 787 AA.

XX

AC ADA36768;

XX

DT 20-NOV-2003 (first entry)

XX

DE Acinetobacter baumannii protein #3929.

XX

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.

XX

OS Acinetobacter baumannii.

XX

PN US6562958-B1.

XX

PD 13-MAY-2003.

XX

PF 04-JUN-1999; 99US-00328352.

XX

PR 09-JUN-1998; 98US-0088701P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Breton G, Bush D;

XX

DR WPI; 2003-576092/54.

XX

DR N-PSDB; ADA32642.

XX

PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

XX

PS Example; SEQ ID NO 8055; 328pp; English.

XX

CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.

```
CC baumannii protein.
XX
SQ Sequence 787 AA;

Query Match 73.2%; Score 41; DB 6; Length 787;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 387 RMADQVAGW 395
|: |||||
|: |||||

RESULT 9
ID AAM92258 standard; protein; 96 AA.
XX
AC AAM92258;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen SEQ ID NO: 1607.
XX
KW Human; digestive system antigen; Gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.
XX
PN WC200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231444P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
```

PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249220P.
 PR 17-NOV-2000; 2000US-0249221P.
 PR 17-NOV-2000; 2000US-0249222P.
 PR 17-NOV-2000; 2000US-0249223P.
 PR 17-NOV-2000; 2000US-0249224P.
 PR 17-NOV-2000; 2000US-0249225P.
 PR 17-NOV-2000; 2000US-0249226P.
 PR 17-NOV-2000; 2000US-0249227P.
 PR 17-NOV-2000; 2000US-0249228P.
 PR 17-NOV-2000; 2000US-0249229P.
 PR 17-NOV-2000; 2000US-0249230P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-502630/55.
 XX N-PSDB; AAK88031.
 DR
 XX
 XX Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases.
 XX
 XX Claim 11; SEQ ID NO 1607; 986pp; English.
 PS
 XX The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a digestive system antigen of
 CC the invention
 XX
 XX Sequence 96 AA;
 SQ
 Query Match 71.4%; Score 40; DB 4; Length 96;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RVRDQVAGW 10
 Db 88 RIRDQLASW 96
 RESULT 10
 ABM91714
 ID ABM91714 standard; protein; 353 AA.
 XX
 XX AC
 XX ABM91714;
 XX
 XX 02-JUN-2005 (first entry)
 XX
 XX DE M. xanthus protein sequence, seq id 10913.
 XX
 XX Transgenic plant; DNA replication; gene regulation; gene expression.
 KW
 XX Myxococcus xanthus.
 OS
 XX
 XX US6833447-B1.
 PN

XX 21-DEC-2004.
 PD
 XX
 XX 10-JUL-2001; 2001US-00902540.
 PF
 XX
 XX 10-JUL-2000; 2000US-0217883P.
 PR
 XX (MONS) MONSANTO TECHNOLOGY LLC.
 XX
 XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
 XX
 XX WPI; 2005-028716/03.
 XX
 DR New substantially purified Myxococcus xanthus nucleic acid molecule
 XX encoding a nitrite reductase, useful for determining gene expression,
 XX identifying mutations in a gene of interest, and for constructing
 PT mutations in a gene of interest.
 PT
 XX Example 2; SEQ ID NO 10913; 25pp; English.
 PS
 XX The invention relates to a substantially purified nucleic acid molecule
 CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
 CC recombinant DNA construct for expression of a nitrite reductase gene in a
 CC plant cell, and a plant cell comprising the recombinant DNA construct.
 CC The nucleic acid is useful for determining gene expression, identifying
 CC mutations in a gene of interest, and for constructing mutations in a gene
 CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
 CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 XX Sequence 353 AA;
 SQ
 Query Match 71.4%; Score 40; DB 9; Length 353;
 Best Local Similarity 60.0%; Pred. No. 40;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GRVRDQVAGW 10
 Db 21 GRIADTVSGW 30
 RESULT 11
 ADN21256
 ID ADN21256 standard; protein; 363 AA.
 XX
 XX AC ADN21256;
 XX
 XX 02-DEC-2004 (first entry)
 XX
 XX Bacterial polypeptide #3909.
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 XX Bacteria.
 OS
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 PD
 XX
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 PA

```

PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 3909; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 363 AA;
SQ
  Query Match          71.4%; Score 40; DB 8; Length 363;
  Best Local Similarity 70.0%; Pred. No. 41;
  Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
  Qy 1 GRVRDQVAGW 10
  Db 243 GEVPDQVVGW 252
  |||||
  |||||

RESULT 12
AAB54291
ID AAB54291 standard; protein; 95 AA.
AC
XX AAB54291;
XX
XX 09-MAR-2001 (first entry)
XX
XX Human pancreatic cancer antigen protein sequence SEQ ID NO:743.
XX
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
XX diagnosis; identification; cytostatic; neuroprotective; nontropic;
XX immunomodulatory; relaxant; contraceptive; gynaecological;
XX antiinflammatory; cardiant; gene therapy; chromosome mapping;
XX linkage analysis; tissue identification; tissue typing; forensic; neural;
XX immune system; muscular; reproductive; gastrointestinal; pulmonary;
XX cardiovascular; renal; proliferative.
XX
XX Homo sapiens.
XX
XX WO200055320-A1.
XX
XX 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005989.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-579444/54.
XX
XX N-PSDB; AAC99056.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition.
XX
XX Claim 11; Page 1188; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
XX AAB54466. The human pancreatic cancer antigens have cytostatic,
XX neuroprotective, nontropic, immunomodulatory, relaxant, contraceptive,
XX gynaecological, cardiant and antiinflammatory activities, and can be used
XX in gene therapy. The polynucleotide and proteins can be used for
XX preventing, treating, or ameliorating a medical condition or in assays
XX for diagnosing a pathological condition or a susceptibility to one in a
XX subject. Binding partners to the proteins and the activity of the
XX proteins can be identified. The pancreatic cancer antigens can be used to
XX detect, treat or prevent pancreatic disorders, especially cancer.
XX Agonists and antagonists to the antigens can be screened for. The
XX pancreatic cancer antigen polynucleotides can be used to design nucleic
XX acid hybridisation probes that can be used in chromosome mapping, linkage
XX analysis, tissue identification and/or typing and a variety of forensic
XX and diagnostic methods. The proteins can be used to generate antibodies
XX which are used to purify, detect and target the polypeptides, including
XX both in vivo and in vitro diagnostic and therapeutic methods. The
XX proteins can be used to treat or prevent neural, immune system, muscular,
XX reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX sequences used in the exemplification of the present invention
XX
SQ Sequence 95 AA;
  Query Match          69.6%; Score 39; DB 3; Length 95;
  Best Local Similarity 60.0%; Pred. No. 16;
  Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
  Qy 1 GRVRDQVAGW 10
  Db 46 GHVRQNIAGW 55
  |||||
  |||||

RESULT 13
AAB47456
ID AAB47456 standard; protein; 558 AA.
AC
XX AAB47456;
XX
XX 13-DEC-2001 (first entry)
XX
XX D-aminoacylase.
XX
XX D-aminoacylase; N-acetyl-D-tryptophan; N-acetyl-D-phenylalanine; PCR;
XX N-acetyl-D-valine; N-acetyl-D-leucine; N-acetyl-D-methionine; primer;
XX probe; synthesis; detection; D-amino acid; substrate specific; amplify;
XX thermal stability; polymerase chain reaction.
XX
XX Hypomyces mycophilus.
XX
XX EPI120465-A1.
XX
XX 01-AUG-2001.
XX
XX 25-JAN-2001; 2001EP-00101739.
XX

```


XX 27-JAN-2000; 2000JP-00019080.
PR 22-MAY-2000; 2000JP-00150578.
XX (DAIL) DAICEL CHEM IND LTD.
XX Mitsuhashi K, Yamamoto H, Matsuyama A, Tokuyama S;
XX WPI; 2001-551332/62.
XX N-PSDB; AAW43261.
XX Novel D-aminoacylase-encoding gene derived from filamentous fungus
PT Hypomyces mycophilus, useful for producing D-tryptophan from N-acetyl-D-
PT tryptophan, useful as medicinal raw material.
XX Claim 1; Page 17-20; 33pp; English.
XX This sequence shows a D-aminoacylase polypeptide. D-aminoacylase has
XX physicochemical properties that include the action of the enzyme on N-
CC acetyl-D-amino acids to produce the corresponding D-amino acids, and
CC substrate specificity, where the enzyme acts on N-acetyl-D-tryptophan, N-
CC acetyl-D-phenylalanine, N-acetyl-D-valine, N-acetyl-L-tryptophan, and N-
CC acetyl-D-methionine, but not on N-acetyl-L-tryptophan, N-acetyl-L-
CC phenylalanine, N-acetyl-L-valine, N-acetyl-L-leucine, or N-acetyl-L-
CC methionine. Fragments of the D-aminoacylase cDNA are useful as primers or
CC as probes for synthesizing or detecting the full length cDNA. D-
CC aminoacylase is useful for producing D-amino acids. The enzyme has
CC substrate specificity, thermal stability and produces D-amino acid
CC efficiently by incubating the fungus derived D-aminoacylase with N-acetyl
CC -D-amino acid under proper conditions. The enzyme has high enzymatic
CC activity for N-acetyl-D-tryptophan and is excellent in industrial
CC applicability. The recombinant polypeptide of D-aminoacylase can be
CC manufactured at a low cost and in large quantities
XX Sequence 558 AA;
SQ
Query Match 69.6%; Score 39; DB 4; Length 558;
Best Local Similarity 62.5%; Pred. No. 99;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 3 VRDOVAGW 10
Db 111 IREQIAGW 118
RESULT 14
ABO62198
ID ABO62198 standard; protein; 358 AA.
XX ABO62198;
XX 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polypeptide seqid 8715.
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX Klebsiella pneumoniae.
XX US6610836-B1.
XX 26-AUG-2003.
XX 27-JAN-2000; 2000US-00489039.
XX 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL, Osborne M;
XX WPI; 2003-895346/82.
DR
XX N-PSDB; ACH95749.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 8715; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX Sequence 358 AA;
SQ
Query Match 67.9%; Score 38; DB 7; Length 358;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GRVRDQVAGW 10
Db 272 GSRURDHVCGW 281
RESULT 15
AAW64365
ID AAW64365 standard; protein; 536 AA.
XX AAW64365;
XX 17-OCT-2003 (revised)
DT 09-NOV-1998 (first entry)
XX Mycobacterium tuberculosis antigen XP14.
XX Tuberculosis; infection; diagnosis; antigen; XP14.
OS Mycobacterium tuberculosis; strain Erdman.
XX WO9816645-A2.
XX 23-APR-1998.
XX 07-OCT-1997; 97WO-US018214.
XX 11-OCT-1996; 96US-00729622.
PR 13-MAR-1997; 97US-00818111.
XX (CORI-) CORIXA CORP.
XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedwick TS, Twardzik DR, Lodes MJ;
XX WPI; 1998-251292/22.
DR N-PSDB; AAW44415.
XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
PT develop products for the detection of M. tuberculosis infection and
PT diagnosis of tuberculosis.
XX Example 3; Page 173-175; 250pp; English.
XX This polypeptide comprises Mycobacterium tuberculosis antigen XP14. Its
CC amino acid sequence was deduced from a DNA molecule (see AAW44415)
CC isolated from a M. tuberculosis strain Erdman genomic DNA expression
CC library using sera from patients having extrapulmonary tuberculosis. XP14
CC shows some sequence similarity to known sequences. The invention relates
CC to compositions and methods for diagnosing tuberculosis. It provides
CC polypeptides (see AAW64291-W64379) comprising an antigenic portion of a
CC soluble M. tuberculosis antigen, or an immunogenic portion of a M.
CC tuberculosis antigen, as well as DNA sequences encoding such

CC polypeptides, recombinant expression vectors and transformed or
 CC transfected host cells. Also claimed are methods and diagnostic kits for
 CC detecting M. tuberculosis infection in a patient using these
 CC polypeptides, antibodies or oligonucleotide probes and primers. (Updated
 CC on 17-OCT-2003 to standardise OS field)

XX
 SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 ||:|||
 Db 335 GRIRDYVRSW 344

RESULT 16
 AAW81732
 ID AAW81732 standard; protein; 536 AA.

XX
 AC AAW81732;

XX
 DT 27-JAN-1999 (first entry)

XX M. tuberculosis immunogenic polypeptide XP14 protein #1.

XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis.

XX Mycobacterium tuberculosis.

XX
 PN WO9816646-A2.

XX
 PD 23-APR-1998.

XX
 PF 07-OCT-1997; 97WO-US018293.

XX
 PR 11-OCT-1996; 96US-00730510.

XX
 PR 13-MAR-1997; 97US-00818112.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
 PI Vedvick TS, Twardzik DR, Lodes MJ;

XX
 DR WPI; 1998-261042/23.

XX
 DR N-PSDB; AAV64524.

XX
 PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
 PT develop products for the detection of M. tuberculosis infection and for
 PT diagnosis, treatment and prevention of tuberculosis.

XX
 PS Example 3c; Page 164-165; 230pp; English.

XX This sequence represents an immunogenic portion of a soluble

CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for
 CC inducing protective immunity against tuberculosis (TB). This sequence can
 CC be formulated into vaccines and/or pharmaceutical compositions for
 CC immunising against M. tuberculosis infection or may be used for the
 CC diagnosis of tuberculosis

XX
 SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 ||:|||
 Db 335 GRIRDYVRSW 344

RESULT 17

AAV39019

ID AAY39019 standard; protein; 536 AA.

XX
 AC AAY39019;

XX
 DT 05-NOV-1999 (first entry)

XX M. tuberculosis recombinant antigen protein XP14.

DE Antigen; diagnosis; detection; infection; antibody; immunisation;
 KW vaccine; immunity.

XX Mycobacterium tuberculosis.

XX
 PN WO9942118-A2.

XX
 PD 26-AUG-1999.

XX
 PF 17-FEB-1999; 99WO-US003265.

XX
 PR 18-FEB-1998; 98US-00024753.

XX
 PR 05-MAY-1998; 98US-00072596.

XX
 PA (CORI-) CORIXA CORP.

XX
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;

PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;

XX
 DR WPI; 1999-527416/44.

XX
 DR N-PSDB; AA219113.

XX
 PT New polypeptide comprising antigenic portions of M. tuberculosis.

XX
 PS Example 3; Page 204-206; 323pp; English.

XX This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against M.
 CC tuberculosis infection. The new detection methods are needed as current
 CC vaccination strategies do not provide 100% immunity

XX
 SQ Sequence 536 AA;

Query Match 67.9%; Score 38; DB 2; Length 536;
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 ||:|||
 Db 335 GRIRDYVRSW 344

RESULT 18

AAV39162

ID AAY39162 standard; protein; 536 AA.

XX
 AC AAY39162;

XX
 DT 05-NOV-1999 (first entry)

XX M. tuberculosis antigen XP14 amino acid sequence.

DE Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;
 KW immune response; skin test.

XX
 OS Mycobacterium tuberculosis.

XX
 PN WO9942076-A2.

XX

```
PD 26-AUG-1999.
XX
XX
XX 17-FEB-1999; 99WO-US003268.
XX
XX 18-FEB-1998; 98US-00025197.
PR 05-MAY-1998; 98US-00072967.
XX
XX (CORI-) CORIXA CORP.
XX
XX Read SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
XX WPI; 1999-527409/44.
XX N-PSDB; AAZ19325.
XX
XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT tests and protective or therapeutic vaccines or compositions.
XX
XX Example 3; Page 159-161; 299pp; English.
XX
XX The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
CC polypeptides fragments, can be used in pharmaceutical compositions or
CC vaccines to generate a protective or therapeutic immune response to M.
CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural
CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
CC to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
CC the present invention
XX
XX Sequence 536 AA;
SQ
Query Match 67.9%; Score 38; DB 2; Length 536;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
DB 335 GRIDYVRSW 344
|||:|:|
|||:|:|

RESULT 19
AAU54158
ID AAU54158 standard; protein; 103 AA.
XX
XX AAU54158;
AC
XX 27-FEB-2002 (first entry)
DT
XX
DE Propionibacterium acnes immunogenic protein #15054.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
PN
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI
```

```
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59563.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 15353; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 103 AA;
SQ
Query Match 66.1%; Score 37; DB 4; Length 103;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
DB 75 GRVTRAIRGW 84
|||:|:|
|||:|:|

RESULT 20
ABM50677
ID ABM50677 standard; protein; 103 AA.
XX
XX ABM50677;
AC
XX 20-OCT-2003 (first entry)
DT
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #15353.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
OS
XX WO2003033515-A1.
PN
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978925.
PR (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
XX WPI; 2003-381789/36.
DR
```

```

DR N-PSDB; ACF64492.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 15353; 1481pp; English.
PS
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 103 AA;
SQ
Query Match 66.1%; Score 37; DB 6; Length 103;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db ||||| : ||
75 GRVETARGW 84

RESULT 21
ABO80351
XX ABO80351 standard; protein; 206 AA.
XX
XX ABO80351;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #12526.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

```

```

XX WPI; 2003-615309/58.
DR N-PSDB; ABD13922.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 29097; 455pp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 206 AA;
SQ
Query Match 66.1%; Score 37; DB 7; Length 206;
Best Local Similarity 70.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db ||||| : |||||
5 GRSRTAVAGW 14

RESULT 22
ABO84669
ID ABO84669 standard; protein; 271 AA.
XX
XX ABO84669;
XX
XX 18-NOV-2004 (first entry)
XX
XX Mouse cancer-associated protein MP20-011.1.
XX
XX Mouse; cancer-associated protein; cytostatic; cancer; leukaemia;
XX lymphoma; CAP.
XX
XX Mus musculus.
XX
XX WO2004074320-A2.
XX
XX 02-SEP-2004.
XX
XX 17-FEB-2004; 2004WO-US0004730.
XX
XX 14-FEB-2003; 2003US-00367094.
XX
XX 14-MAR-2003; 2003US-00388838.
XX
XX 15-APR-2003; 2003US-00417375.
XX
XX 13-JUN-2003; 2003US-00461862.
XX
XX 15-SEP-2003; 2003US-00663431.
XX
XX 15-DEC-2003; 2003US-00737318.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Morris DW, Malandro MS;
XX
XX WPI; 2004-652914/63.
DR N-PSDB; ABD32952.
XX

```

PT New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX disclosure; seqid 697; 310pp; English.

PS The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP protein sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 271 AA;

Query Match 66.1%; Score 37; DB 8; Length 271;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GRVRDQVAGW 10
|::| |
Db 202 GKIKDSNAGW 211

RESULT 23
AAG92549
ID AAG92549 standard; protein; 297 AA.

XX AAG92549;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 6303.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH67768.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 6303; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids.
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX Sequence 297 AA;

Query Match 66.1%; Score 37; DB 4; Length 297;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 RVRDQVAGW 10

||| |

Db 74 RVRIQVTGW 82

RESULT 24

ADT57937

ID ADT57937 standard; protein; 423 AA.

XX ADT57937;

XX 13-JAN-2005 (first entry)

XX Plant polypeptide, SEQ ID 8014.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.

XX Viridiplantae.

XX US2004216190-A1.

XX 28-OCT-2004.

XX 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

XX 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

XX WPI; 2004-757369/74.

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.

XX
PS Claim 2; SEQ ID NO 8014; 14pp; English.

XX
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX
SQ Sequence 423 AA;

Query Match 66.1%; Score 37; DB 8; Length 423;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
| : ||| : |
Db 366 GMIRDQILSW 375

RESULT 25

AAG91159
ID AAG91159 standard; protein; 605 AA.

XX
AC AAG91159;

DT 26-SEP-2001 (first entry)

XX
DE C glutamicum protein fragment SEQ ID NO: 4913.

XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.

XX
OS Corynebacterium glutamicum.

XX
FN EP1108790-A2.

XX
XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR N-PSDB; AAH66378.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 4913; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX
SQ Sequence 605 AA;

Query Match 66.1%; Score 37; DB 4; Length 605;
Best Local Similarity 70.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
| : ||| : |
Db 402 GRDTDQVAAM 411

RESULT 26

ADS29923
ID ADS29923 standard; protein; 746 AA.

XX
AC ADS29923;

DT 02-DEC-2004 (first entry)

XX
DE Bacterial polypeptide #18956.

XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX
OS Bacteria.

XX
PN US2003233675-A1.

XX
PD 18-DEC-2003.

XX
PF 20-FEB-2003; 2003US-00369493.

XX
PR 21-FEB-2002; 2002US-0360039P.

XX

Mon May 15 11:35:31 2006

```
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 18956; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 746 AA;
XX
XX Query Match 66.1%; Score 37; DB 8; Length 746;
XX Best Local Similarity 66.7%; Pred. No. 3.2e+02;
XX Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 2 RVRDQVAGN 10
XX |: ||| ||
XX Db 344 RLADQVTGW 352
XX
XX RESULT 27
XX AAU64755
XX ID AAU64755 standard; protein; 66 AA.
XX
XX AC AAU64755;
XX
XX 27-FEB-2002 (first entry)
XX
XX DE Propionibacterium acnes immunogenic protein #25651.
XX
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO200181581-A2.
XX
XX PD 01-NOV-2001.
XX
XX PF
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX
XX 02-JUN-2000; 2000US-0208841P.
XX
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX
XX N-PSDB; AAS9650.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 25950; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 66 AA;
XX
XX Query Match 64.3%; Score 36; DB 4; Length 66;
XX Best Local Similarity 60.0%; Pred. No. 42;
XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 GRVRDQVAGN 10
XX ||| ||: ||
XX Db 9 GRPRDRRSGW 18
XX
XX RESULT 28
XX ABM61274
XX ID ABM61274 standard; protein; 66 AA.
XX
XX AC ABM61274;
XX
XX 20-OCT-2003 (first entry)
XX
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #25950.
XX
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX OS Propionibacterium acnes.
XX
XX PN WO2003033515-A1.
XX
XX PD 24-APR-2003.
XX
XX PF 11-OCT-2002; 2002WO-US032727.
XX
```

```
PR 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
PA
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64579.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 25950; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 66 AA;
Query Match 64.3%; Score 36; DB 6; Length 66;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
Db ||||: |||
9 GRPRDRSGW 18
RESULT 29
AAB41784
ID AAB41784 standard; protein; 159 AA.
XX
XX AAB41784;
XX
XX 08-FEB-2001 (first entry)
XX Human ORFX ORF1548 polypeptide sequence SEQ ID NO:3096.
XX
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX vulnery; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
```

```
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection; asthma;
KW severe combined immunodeficiency; malaria; autoimmune disorder; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
XX 02-APR-1999; 99US-0127636P.
XX 05-APR-1999; 99US-0127728P.
XX 30-MAR-2000; 2000US-00540763.
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX N-PSDB; AAC75993.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 11; Page 2311; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
XX Sequence 159 AA;
Query Match 64.3%; Score 36; DB 3; Length 159;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
Db ||||: |||
71 GRVDAVDGW 80
RESULT 30
ABM97186
ID ABM97186 standard; protein; 290 AA.
XX
XX ABM97186;
XX
XX 02-JUN-2005 (first entry)
XX
```


XX M. xanthus protein sequence, seq id 16385.
DE Transgenic plant; DNA replication; gene regulation; gene expression.
XX Myxococcus xanthus.
XX US6833447-B1.
XX 21-DEC-2004.
XX 10-JUL-2001; 2001US-00902540.
XX 10-JUL-2000; 2000US-0217883P.
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX Example 2; SEQ ID NO 16385; 25pp; English.
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX Sequence 290 AA;
Query Match 64.3%; Score 36; DB 9; Length 290;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 RVRDQVAGW 10
DB 191 RVRDEVRAW 199
RESULT 31
ABB84757
ID ABB84757 standard; protein; 336 AA.
XX ABB84757;
XX 16-MAY-2002 (first entry)
XX DNA polymerase III holoenzyme delta subunit related protein SEQ ID:102.
XX DNA polymerase III holoenzyme delta subunit; DNA polymerase III; DnaX;
KW hola; holB; antibacterial; tuberculostatic; antileprotic;
KW bacterial infection; DNA replication modulation.
XX Streptomyces coelicolor.
XX WO200206532-A1.
XX 24-JAN-2002.
XX 16-JUL-2001; 2001WO-US022395.
XX 14-JUL-2000; 2000US-0218246P.
XX Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit from

PR 28-MAR-2001; 2001US-00818780.
XX (REPL-) REPLIDYNE INC.
PA Bullard JJ, Janjic N, Mchenry CS;
XX WPI; 2002-164785/21.
XX N-PSDB; ABL88001.
XX Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit from
PT bacteria, useful for screening agents that modulate the subunit activity
PT which is useful in the treatment of bacterial infections e.g. S. pyogenes
PT and S. aureus.
XX Claim 19; Page 340-341; 500pp; English.
XX The present invention describes nucleic acid sequences encoding a DNA
CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
CC antibacterial; tuberculostatic; antileprotic. Methods from the present
CC invention can be used for screening for bacterial DNA polymerase
CC holoenzyme delta subunit proteins and agents that modulate their
CC activity. The agents are useful in the treatment of bacterial infections
CC e.g. S. pyogenes, S. aureus, S. pneumoniae, Mycoplasma, Yersinia,
CC Corynebacterium, Salmonella, Mycobacterium tuberculosis or M. leprae. The
CC invention provides a convenient means of identifying compounds which
CC modulate DNA replication in bacteria and therefore provide antibacterial
CC targets, and which are also useful for amplification of DNA. ABL87935 to
CC ABL88071 and ABB84724 to ABB84816 represent sequences used in the
CC exemplification of the present invention
XX Sequence 336 AA;
Query Match 64.3%; Score 36; DB 5; Length 336;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 RVRDQVAGW 10
DB 281 RVRQMRGW 289
RESULT 32
ABB84815
ID ABB84815 standard; protein; 336 AA.
XX ABB84815;
XX 16-MAY-2002 (first entry)
XX DNA polymerase III holoenzyme delta subunit related protein SEQ ID:229.
XX DNA polymerase III holoenzyme delta subunit; DNA polymerase III; DnaX;
KW hola; holB; antibacterial; tuberculostatic; antileprotic;
KW bacterial infection; DNA replication modulation.
XX Streptomyces coelicolor.
XX WO200206532-A1.
XX 24-JAN-2002.
XX 16-JUL-2001; 2001WO-US022395.
XX 14-JUL-2000; 2000US-0218246P.
XX 28-MAR-2001; 2001US-00818780.
XX (REPL-) REPLIDYNE INC.
XX Bullard JJ, Janjic N, Mchenry CS;
XX WPI; 2002-164785/21.
XX Nucleic acid encoding a DNA polymerase III holoenzyme delta subunit from

PT bacteria, useful for screening agents that modulate the subunit activity
 PT which is useful in the treatment of bacterial infections e.g. *S. pyogenes*
 PT and *S. aureus*.

XX Claim 19; Fig 7F; 500pp; English.

XX CC The present invention describes nucleic acid sequences encoding a DNA
 CC polymerase III holoenzyme delta subunit (I) from bacteria. (I) has
 CC antibacterial, tuberculostatic; antileprotic. Methods from the present
 CC invention can be used for screening for bacterial DNA polymerase
 CC holoenzyme delta subunit proteins and agents that modulate their
 CC activity. The agents are useful in the treatment of bacterial infections
 CC e.g. *S. pyogenes*, *S. aureus*, *S. pneumoniae*, *Mycoplasma*, *Yersinia*,
 CC *Corynebacterium*, *Salmonella*, *Mycobacterium tuberculosis* or *M. leprae*. The
 CC invention provides a convenient means of identifying compounds which
 CC modulate DNA replication in bacteria and therefore provide antibacterial
 CC targets, and which are also useful for amplification of DNA. ABL87935 to
 CC ABL88071 and ABB84724 to ABB84816 represent sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 336 AA;

Query Match 64.3%; Score 36; DB 5; Length 336;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10

Db 281 RVRQMRGW 289

RESULT 33

ABG01473

ID ABG01473 standard; protein; 343 AA.

XX AC ABG01473;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #1464.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS65660.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 31832; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological actions. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 343 AA;

Query Match 64.3%; Score 36; DB 4; Length 343;

Best Local Similarity 50.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDOVAGW 10

Db 89 GKLRPEAGW 98

RESULT 34

ABU41540

ID ABU41540 standard; protein; 355 AA.

XX AC ABU41540;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by prokaryotic essential gene #27067.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Pseudomonas syringae.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA45410.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 69464; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 355 AA;

Query Match 64.3%; Score 36; DB 6; Length 355;
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GRVRDQVAGW 10
 |||:|:|
 Db 168 GYVQRQAGW 177

RESULT 35
 ADH18901
 ID ADH18901 standard; protein; 385 AA.
 XX
 AC ADH18901;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human cell adhesion and extracellular matrix CADECM-28 protein - SEQ 28.
 XX
 KW human; cell adhesion and extracellular matrix; CADECM; cytostatic;
 KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
 KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
 KW antiinflammatory; ophthalmological; antithyroid; antiarthritic;
 KW antibacterial; virucide; protozoacide; antiparasitic; fungicide;
 KW anorectic; cardiant; hypotensive; antiinfertility; hepatotropic;
 KW autoimmune; inflammatory; AIDS; allergy; atopic dermatitis; arthritis;
 KW thyroiditis; bacterial infection; viral; parasitic; protozoal; fungal;
 KW metabolic; obesity; reproductive; infertility; neurological;
 KW Parkinson's disease; Alzheimer's; stroke; epilepsy; cardiovascular;
 KW myocardial infarction; hypertension; eye; cell proliferative; cancer;
 KW atherosclerosis; hepatitis; SNP; single nucleotide polymorphism.
 XX
 OS Homo sapiens.
 XX
 PN WO2003094843-A2.
 XX
 PD 20-NOV-2003.
 XX
 PF 06-MAY-2003; 2003WO-US014076.
 XX
 PR 10-MAY-2002; 2002US-0379840P.
 PR 17-MAY-2002; 2002US-0381291P.
 PR 24-MAY-2002; 2002US-0383183P.

PR 05-JUL-2002; 2002US-0394146P.
 XX (INCY-) INCYTE CORP.
 PA
 XX Khare R, Elliott VS, Marquis JP, Ramkumar J, Chawla NK, Mason PM;
 PI Hafalia AJA, Swarnakar A, Jin P, Becha SD, Kable AE, Tran UK;
 PI Baughn MR, Burford N, Graul RC, Emerling BW, Sprague WW, Griffin JA,
 PI Ison CH;
 XX WPI; 2004-0111995/01.
 DR N-PSDB; ADH18932.
 XX
 PT New human cell adhesion and extracellular matrix proteins and
 PT polynucleotides, useful for diagnosing, preventing or treating diseases
 PT or conditions associated with aberrant protein expression, e.g. cancer,
 PT AIDS or stroke.
 XX
 PS Claim 1; SEQ ID NO 28; 308pp; English.
 XX
 CC The invention relates to a novel isolated human cell adhesion and
 CC extracellular matrix (CADECM) polypeptide. The polypeptide of the
 CC invention demonstrates cytostatic, antiarteriosclerotic, anti-HIV,
 CC antiallergic, cerebroprotective, antiparkinsonian, anticonvulsant,
 CC nootropic, neuroprotective, antiinflammatory, ophthalmological,
 CC antithyroid, antiarthritic, antiparasitic, antiparasitic, fungicide,
 CC antiparasitic, fungicide, anorectic, cardiant, hypotensive,
 CC antiinfertility and hepatotropic activities. The polypeptide may be
 CC useful in diagnosing, preventing or treating diseases or conditions such
 CC as autoimmune or inflammatory disorders including AIDS, allergies, atopic
 CC dermatitis, arthritis or thyroiditis, infections particularly bacterial,
 CC viral, parasitic, protozoal or fungal, metabolic disorders such as
 CC obesity, reproductive disorders e.g. infertility, neurological disorders
 CC including Parkinson's disease, Alzheimer's disease, stroke or epilepsy,
 CC cardiovascular disorders e.g. myocardial infarction or hypertension, eye
 CC disorders or cell proliferative diseases including cancer,
 CC atherosclerosis and hepatitis. The current sequence is that of the human
 CC CADECM protein of the invention.
 XX
 SQ Sequence 385 AA;
 Query Match 64.3%; Score 36; DB 8; Length 385;
 Best Local Similarity 70.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 GRVRDQVAGW 10
 |||:|:|
 Db 348 GRVVDVADGW 357
 XX
 DE Human cell adhesion and extracellular matrix CADECM-28 protein - SEQ 28.
 XX
 AC ABO81727;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #13902.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

CC The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC
SQ Sequence 423 AA;

Query Match 64.3%; Score 36; DB 5; Length 423;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
||| ||| |||
Db 179 GRVVDVAGW 188

RESULT 39
AAO30830
ID AAO30830 standard; protein; 436 AA.

XX AC AAO30830;
XX AC

XX DT 22-SEP-2003 (first entry)
XX DE Human cell adhesion and extracellular matrix protein (CADECM)-20.
XX KW Human; cell adhesion and extracellular matrix protein; immune disorder;
XX KW CADECM; cancer; gene therapy.
XX OS Homo sapiens.
XX PN WO2003047526-A2.
XX PD 12-JUN-2003.
XX PF 26-NOV-2002; 2002WO-US038437.
XX PR 30-NOV-2001; 2001US-0334343P.
XX PR 07-DEC-2001; 2001US-0340278P.
XX PR 04-JAN-2002; 2002US-0345069P.
XX PR 25-JAN-2002; 2002US-0351352P.
XX PR 14-FEB-2002; 2002US-0357168P.
XX PR 29-MAR-2002; 2002US-0369128P.
XX PR 05-APR-2002; 2002US-0370802P.
XX PA (INCYTE) INCYTE GENOMICS INC.
XX PI Baughn MR, Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD;
XX PI Deleage AM, Elliott VS, Gandhi AR, Gierzen KJ, Gorvad AE;
XX PI Griffin JA, Ho A, Jin P, Kable AE, Lal PG, Lee EA, Lee S, Lee SY;
XX PI Marquis JP, Lehr-Mason PM, Ramkumar J, Richardson TW, Sprague MW;
XX PI Swarnakar A, Tang TV, Tran B, Tran UK, Chawla NK, Warren BA, Xu Y;
XX PI Yue H, Zheng W;
XX WPI; 2003-513695/48.
XX DR N-PSDB; AAL62032.

XX PT New human cell adhesion and extracellular matrix proteins (CADECM) polypeptide, useful for preparing a composition for treating a disease associated with decreased expression or overexpression of CADECM e.g.,

PT cancer.

XX Claim 1; Page 300-301; 374pp; English.

XX CC The invention relates to human cell adhesion and extracellular matrix proteins (CADECM) and nucleic acid molecules encoding such proteins. CADECM proteins are useful for preparing a composition for diagnosing or treating a disease or condition associated with decreased expression or overexpression of functional CADECM e.g., immune disorders or cancer. The invention is also useful in gene therapy. The present sequence is human CADECM protein

XX SQ Sequence 436 AA;

Query Match 64.3%; Score 36; DB 7; Length 436;
Best Local Similarity 70.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
||| ||| |||
Db 348 GRVVDVAGW 357

RESULT 40
ABB90892
ID ABB90892 standard; protein; 479 AA.

XX AC ABB90892;
XX AC
XX DT 31-MAY-2002 (first entry)
XX DE Herbicidally active polypeptide SEQ ID NO 103.
XX KW Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX PN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EP009892.
XX PR 28-AUG-2001; 2001WO-EP009892.
XX PA (FARB) BAYER AG.
XX PI Tietjen K, Weidler M;
XX WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides

XX SQ Sequence 479 AA;

Query Match 64.3%; Score 36; DB 5; Length 479;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

PS Claim 5; SEQ ID NO 103; 261pp + Sequence Listing; English.

XX CC The invention relates to identifying target proteins (ABB90750-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides

XX SQ Sequence 479 AA;

Query Match 64.3%; Score 36; DB 5; Length 479;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
 DB 269 RDQVGVW 275

RESULT 41
 ID ADT56368 standard; protein; 479 AA.
 AC ADT56368;
 XX
 XX 13-JAN-2005 (first entry)
 XX
 DE Plant polypeptide, SEQ ID 6445.
 XX
 KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
 KW disease resistance; galactomannan production; plant growth regulator;
 KW heat tolerance; herbicide tolerance; lignin production;
 KW extreme osmotic condition tolerance; pathogens resistance;
 KW pest resistance; yield improvement; seed oil yield; seed protein yield.
 XX
 OS Viridiplantae.
 XX
 XX US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 XX 18-DEC-2003; 2003US-00739930.
 PF
 XX 28-APR-2003; 2003US-00424599.
 PR
 XX 28-APR-2003; 2003US-00425115.
 XX
 PA (KOVA/) KOVALIC D K.
 XX
 XX Kovalic DK;
 PI
 XX WPI; 2004-757369/74.
 DR
 XX New recombinant DNA constructs useful in the field of biochemistry and
 PT genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX
 XX Claim 2; SEQ ID NO 6445; 14pp; English.
 PS
 XX The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of
 CC producing a plant having an improved property, comprising transforming a
 CC plant with a recombinant DNA construct comprising a promoter region
 CC functional in a plant cell operably joined to a polynucleotide encoding a
 CC polypeptide associated with the property, and growing the transformed
 CC plant. The property is selected from improving plant cold tolerance, for
 CC manipulating growth rate in plant cells by modification of the cell cycle
 CC pathway, for improving plant drought tolerance, for providing increased
 CC resistance to plant disease, for galactomannan production, for production
 CC of plant growth regulators, for improving plant heat tolerance, for
 CC improving plant tolerance to herbicides, for increasing the rate of
 CC homologous recombination in plants, for lignin production, for improving
 CC plant tolerance to extreme osmotic conditions, for improving plant
 CC tolerance to pathogens or pests, for yield improvement by modification of
 CC photosynthesis, for modifying seed oil yield and/or content, for
 CC modifying seed protein yield and/or content, for yield improvement by
 CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
 CC and for yield improvement by providing improved plant growth and
 CC development under at least one stress condition. The polynucleotide may
 CC also encode a plant transcription factor. The methods and compositions of
 CC the present invention are useful in the field of biochemistry and
 CC genetics, in particular for producing transgenic plants with improved
 CC biological characteristics such as increased yield, improved nitrogen

CC flow, increasing plant tolerance to cold or heat, improving plant
 CC tolerance to extreme osmotic and drought conditions, and improving plant
 CC tolerance to plant pests or pathogens. They can also be used in physical
 CC arrays of molecules, plant breeding markers, computer-based storage and
 CC analysis systems. The present sequence is one of the 5544 plant protein
 CC sequences of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
 XX
 SQ Sequence 479 AA;
 Query Match 64.3%; Score 36; DB 8; Length 479;
 Best Local Similarity 71.4%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
 DB 269 RDQVGVW 275

RESULT 42
 ID ADV42198 standard; protein; 479 AA.
 AC ADV42198;
 XX
 XX 24-FEB-2005 (first entry)
 DT
 XX
 DE Glycosyltransferase UGT71C4 amino acid sequence.
 XX
 KW Glycosyltransferase; protein engineering; enzyme engineering; enzyme.
 XX
 OS Unidentified.
 XX
 PN WO2004106508-A2.
 XX
 PD 09-DEC-2004.
 XX
 XX 24-MAY-2004; 2004WO-GB002237.
 PF
 XX 27-MAY-2003; 2003GB-00012042.
 PR
 XX 28-JUN-2003; 2003GB-00015183.
 XX
 XX (UYYO-) UNIV YORK.
 PA
 XX Lim EK, Bowles D;
 PI
 XX WPI; 2005-021284/02.
 DR
 XX N-PSDB; ADV42197.
 DR
 XX New reaction vessel comprising a genetically modified cell and a nutrient
 PT medium for supporting the growth of the cell that includes a substrate
 PT for glycosyltransferase, useful for producing recombinant proteins, e.g.
 PT proteases or amylase.
 XX
 XX Claim 6; Fig 3b; 72pp; English.
 PS
 XX The invention comprises a reaction vessel (e.g. a bioreactor) comprising
 CC a genetically modified cell, wherein the cell is modified by transfection
 CC or transformation with a nucleic acid that encodes a glycosyltransferase
 CC enzyme. The reaction vessel of the invention is useful for the production
 CC of molecules (e.g. recombinant proteins) on an industrial scale. The
 CC present amino acid sequence represents a glycosyltransferase enzyme of
 CC the invention.
 XX
 SQ Sequence 479 AA;
 Query Match 64.3%; Score 36; DB 9; Length 479;
 Best Local Similarity 71.4%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDQVAGW 10

CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 3; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 270 GRVVDVAGW 279
|||||

RESULT 44
AAM93202
ID AAM93202 standard; protein; 517 AA.
XX
AC AAM93202;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 2591.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EPI130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94110.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2591; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
SQ Sequence 517 AA;

Query Match 64.3%; Score 36; DB 4; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
Db 270 GRVVDVAGW 279
|||||

Db 269 RDQIVGW 275
|||||

RESULT 43
AAB41790
ID AAB41790 standard; protein; 517 AA.
XX
AC AAB41790;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1554 polypeptide sequence SEQ ID NO:3108.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127638P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR N-PSDB; AAC75999.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 11; Page 2331-2332; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antirheumatic; immunosuppressant; immunostimulant;
CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and

```

RESULT 45
ID AAM93206 standard; protein; 517 AA.
XX AC AAM93206;
XX DT 06-NOV-2001 (first entry)
XX DB Human polypeptide, SEQ ID NO: 2599.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2001-524255/58.
XX DR N-PSDB; AAK94114.
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Claim 8; SEQ ID NO 2599; 1380pp + Sequence Listing; English.
XX CC The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a polypeptide encoded by a full length
XX CC human cDNA of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in CD-ROM
XX CC format directly from BPO
XX SQ Sequence 517 AA;
    Query Match 64.3%; Score 36; DB 4; Length 517;
    Best Local Similarity 70.0%; Pred. No. 3.4e+02;
    Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVVDQVAGW 10
DB 270 GRVVDQVAGW 279
    |||||
    |||||

RESULT 46
ADL30558
ID ADL30558 standard; protein; 517 AA.
XX AC ADL30558;
XX DT 20-MAY-2004 (first entry)
XX DE Human protein encoded by a full length cDNA clone SeqID 2591.
XX KW Human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX OS Homo sapiens.
XX PN EP1396543-A2.
XX PD 10-MAR-2004.
XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX KW Human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX OS Homo sapiens.
XX PN EP1396543-A2.
XX PD 10-MAR-2004.
XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.

KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method.
XX Homo sapiens.
XX PN EP1396543-A2.
XX PD 10-MAR-2004.
XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2004-204755/20.
XX DR N-PSDB; ADL30557.
XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
XX PT length human cDNAs.
XX PS Example 1; SEQ ID NO 2591; 1340pp; English.
XX CC This invention relates to a novel primers useful for synthesising full
XX CC length cDNA molecules that encode human proteins. Specifically, it refers
XX CC to secretory or membrane proteins that are potential therapeutic agents/
XX CC target molecules in the field of medicine, and in particular genes
XX CC encoding proteins that are associated with signal transduction,
XX CC glycoproteins and transcription. The present invention describes a method
XX CC for efficiently cloning a full length human cDNA from both the 5' and 3'
XX CC ends using the oligo-capping method. This polypeptide sequence is a full
XX CC length human protein of the invention.
XX SQ Sequence 517 AA;
    Query Match 64.3%; Score 36; DB 8; Length 517;
    Best Local Similarity 70.0%; Pred. No. 3.4e+02;
    Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVVDQVAGW 10
DB 270 GRVVDQVAGW 279
    |||||
    |||||

RESULT 47
ADL30566
ID ADL30566 standard; protein; 517 AA.
XX AC ADL30566;
XX DT 20-MAY-2004 (first entry)
XX DE Human protein encoded by a full length cDNA clone SeqID 2599.
XX KW human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX OS Homo sapiens.
XX PN EP1396543-A2.
XX PD 10-MAR-2004.
XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.

```



```

PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR N-PSDB; ADL30565.
XX
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 1; SEQ ID NO 2599; 1340pp; English.
XX
XX This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polypeptide sequence is a full
CC length human protein of the invention.
XX
SQ Sequence 517 AA;
Query Match 64.3%; Score 36; DB 8; Length 517;
Best Local Similarity 70.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GRVRDQVAGW 10
Db |||||
270 GRVVDAVDGW 279

RESULT 48
ADH18902
ID ADH18902 standard; protein; 527 AA.
XX
AC ADH18902;
XX
XX 11-MAR-2004 (first entry)
DT
DE Human cell adhesion and extracellular matrix CADECM-29 protein - SEQ 29.
XX
XX human; cell adhesion and extracellular matrix; CADECM; cytostatic;
KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
KW antiinflammatory; ophthalmological; antithyroid; antiarthritic;
KW antibacterial; virucide; protozoacide; antiparasitic; fungicide;
KW anorectic; cardiant; hypotensive; antiinfertility; hepatotropic;
KW autoimmune; inflammatory; AIDS; allergy; atopic dermatitis; arthritis;
KW thyroiditis; bacterial infection; viral; parasitic; protozoal; fungal;
KW metabolic; obesity; reproductive; infertility; neurological;
KW Parkinson's disease; Alzheimer's; stroke; epilepsy; cardiovascular;
KW myocardial infarction; hypertension; eye; cell proliferative; cancer;
KW atherosclerosis; hepatitis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX WO2003094843-A2.
PN
XX
PD 20-NOV-2003.
XX
XX 06-MAY-2003; 2003WO-US014076.
PF
XX 10-MAY-2002; 2002US-0379840P.
PR 17-MAY-2002; 2002US-0381291P.
PR 24-MAY-2002; 2002US-0383183P.
PR 05-JUL-2002; 2002US-0394146P.
XX

```

```

PA (INCY-) INCYTE CORP.
XX
XX Khare R, Elliott VS, Marquis JP, Ramkumar J, Chawla NK, Mason PM;
PI Hafalia AJA, Swarnakar A, Jin P, Becha SD, Kable AE, Tran UK;
PI Baughn MR, Burford N, Graul RC, Emerling BM, Sprague WM, Griffin JA;
PI Ison CH;
XX
XX WPI; 2004-011995/01.
DR N-PSDB; ADH18933.
XX
XX New human cell adhesion and extracellular matrix proteins and
PT polynucleotides, useful for diagnosing, preventing or treating diseases
PT or conditions associated with aberrant protein expression, e.g. cancer,
PT AIDS or stroke.
XX
PS Claim 1; SEQ ID NO 29; 308pp; English.
XX
XX The invention relates to a novel isolated human cell adhesion and
CC extracellular matrix (CADECM) polypeptide. The polypeptide of the
CC invention demonstrates cytostatic, antiarteriosclerotic, anti-HIV,
CC antiallergic, cerebroprotective, antiparkinsonian, anticonvulsant,
CC nootropic, neuroprotective, antiinflammatory, ophthalmological,
CC antithyroid, antiarthritic, antibacterial, virucide, protozoacide,
CC antiparasitic, fungicide, anorectic, cardiant, hypotensive,
CC antiinfertility and hepatotropic activities. The polypeptide may be
CC useful in diagnosing and inflammatory disorders including AIDS, allergies, atopic
CC as autoimmune or inflammatory disorders including AIDS, allergies, atopic
CC dermatitis, arthritis or thyroiditis, infections particularly bacterial,
CC viral, parasitic, protozoal or fungal, metabolic disorders such as
CC obesity, reproductive disorders e.g. infertility, neurological disorders
CC including Parkinson's disease, Alzheimer's disease, stroke or epilepsy,
CC cardiovascular disorders e.g. myocardial infarction or hypertension, eye
CC disorders or cell proliferative diseases including cancer,
CC atherosclerosis and hepatitis. The current sequence is that of the human
CC CADECM protein of the invention.
XX
SQ Sequence 527 AA;
Query Match 64.3%; Score 36; DB 8; Length 527;
Best Local Similarity 70.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GRVRDQVAGW 10
Db |||||
348 GRVVDAVDGW 357

RESULT 49
ADQ21291
ID ADQ21291 standard; protein; 560 AA.
XX
AC ADQ21291;
XX
XX 26-AUG-2004 (first entry)
DT
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4111.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
KW
XX Homo sapiens.
XX
XX WO2004048938-A2.
PN
XX
PD 10-JUN-2004.
XX
XX 26-NOV-2003; 2003WO-US038193.
PF
XX 26-NOV-2002; 2002US-0429739P.
PR
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX

```

DR WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 4111; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX
SQ Sequence 560 AA;
Query Match 64.3%; Score 36; DB 8; Length 560;
Best Local Similarity 70.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
DB 347 GRVVDVADGW 356
|||||
347 GRVVDVADGW 356
RESULT 50
AAG81257
ID AAG81257 standard; protein; 578 AA.
XX
XX AAG81257;
XX
DT 10-SEP-2001 (first entry)
XX
XX Human AFP protein sequence SEQ ID NO:32.
XX
XX Human; secreted protein; secretion; bacterial cell; fungal cell;
XX eukaryotic cell; fusion protein; maltose binding protein;
XX immunoglobulin constant region; polyhistidine tag.
XX
XX Homo sapiens.
XX
XX WO200129221-A2.
XX
XX 26-APR-2001.
XX
XX 20-OCT-2000; 2000WO-US029052.
XX
XX 20-OCT-1999; 99US-0160712P.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Conklin DC, Yee DP;
XX
XX WPI; 2001-300340/31.
DR N-PSDB; AAH52108.
XX
XX Isolated polypeptide for directing secretion of proteins of interest from
PT a host cell including, e.g. bacteria, includes contiguous amino acid
PT residues of polypeptide with specified amino acids.
XX
XX Claim 1; Page 121-122; 617pp; English.
XX
XX AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
CC to AAG81453. The secreted proteins can be used for directing the
CC

CC secretion of proteins of interest from a host cell including bacteria,
CC fungal cells, and cultured higher eukaryotic cells. The present invention
CC also describes fusion proteins, where a secreted protein of the invention
CC is operably linked via a peptide bond or peptide linker to a second
CC protein selected from the group consisting of maltose binding protein, an
CC immunoglobulin constant region, a polyhistidine tag and a peptide given
CC in AAG81453
XX
XX Sequence 578 AA;
Query Match 64.3%; Score 36; DB 4; Length 578;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
DB 348 GRVVDVADGW 357
|||||
348 GRVVDVADGW 357
Search completed: May 12, 2006, 10:34:53
Job time : 115.75 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-35
Sequence: 56
1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

PIR 80: *
1: Pirl: *
2: Pirl: *
3: Pirl: *
4: Pirl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	73.2	780	1 S76487	probable copper-tr
2	39	69.6	320	2 E87684	ATP phosphoribosyl
3	39	69.6	360	2 AD1929	hypothetical prote
4	39	69.6	1216	2 T34101	hypothetical prote
5	38	67.9	300	2 B70725	probable linB prot
6	37	66.1	63	2 AB2809	conserved hypothet
7	37	66.1	181	2 C84233	hypothetical prote
8	37	66.1	218	1 S02828	albicidin resistan
9	37	66.1	421	2 C95880	probable protein l
10	37	66.1	627	2 AI2146	excinuclease ABC c
11	37	66.1	747	1 S36741	probable copper-tr
12	37	66.1	753	2 AH2543	cation-transportin
13	37	66.1	753	2 AE2009	cation-transportin
14	37	66.1	1752	2 A45407	collagen alpha 3(I
15	36	64.3	67	2 G82724	hypothetical prote
16	36	64.3	283	2 C75353	replication protei
17	36	64.3	338	2 C75353	ribosomal large su
18	36	64.3	356	2 E83416	hypothetical prote
19	36	64.3	415	2 F70686	hypothetical prote
20	36	64.3	479	2 G86207	hypothetical prote
21	36	64.3	723	2 I39066	N-methyl-D-asparta
22	36	64.3	951	2 G82965	conserved hypothet
23	36	64.3	1260	4 GNLR1	retrovirus-related
24	36	64.3	2517	2 S58380	probable RNA-direc
25	35	62.5	113	2 G82245	transcription regu
26	35	62.5	130	2 B69047	arsenate reductase
27	35	62.5	150	2 A41840	hypothetical prote
28	35	62.5	257	2 AD2083	phosphonate ABC tr
29	35	62.5	300	2 S76855	hypothetical prote

30	62.5	302	2 B87128	conserved hypothet
31	62.5	370	2 G83219	probable dihydroli
32	62.5	428	2 S38461	ubiquinol-cytochro
33	62.5	482	2 G84823	probable anthocyana
34	62.5	733	2 E95335	probable cation tr
35	62.5	891	2 H75507	hypothetical prote
36	62.5	1104	2 C72409	reverse gyrase - T
37	62.5	1273	2 E72611	probable ATP-depen
38	60.7	154	2 AC3434	calf thymus ribonu
39	60.7	186	2 C72782	hypothetical prote
40	60.7	193	2 D83620	hypothetical prote
41	60.7	213	2 G87457	guanylate kinase l
42	60.7	225	2 C84229	hypothetical prote
43	60.7	252	2 F87259	hypothetical prote
44	60.7	271	2 AI3628	peptidylprolyl iso
45	60.7	337	2 T16323	hypothetical prote
46	60.7	343	2 G71499	probable D-Ala-D-A
47	60.7	434	2 G95297	probable hydrolase
48	60.7	457	2 C85203	argininosuccinate
49	60.7	457	2 F86087	argininosuccinate
50	60.7	457	2 A98240	probable argininos
51	60.7	457	2 A10477	argininosuccinate
52	60.7	458	2 AH0935	argininosuccinate
53	60.7	488	1 H64537	cytochrome-c oxida
54	60.7	488	2 G71969	cytochrome-c oxida
55	60.7	512	2 T30924	hypothetical prote
56	60.7	545	1 RRPBPM	RNA-directed RNA p
57	60.7	548	2 A39094	glucan endo-1,3-be
58	60.7	569	2 J41777	Pal2 protein - yea
59	60.7	901	2 AG0996	Malt regulatory pr
60	60.7	1115	2 T29012	hypothetical prote
61	60.7	1166	2 F90247	reverse gyrase (to
62	60.7	1166	2 T29099	reverse gyrase - S
63	60.7	1177	2 G70399	reverse gyrase - A
64	60.7	1214	2 H75034	reverse gyrase (to
65	60.7	1221	2 T18550	reverse gyrase cha
66	60.7	1222	2 G72614	probable reverse g
67	60.7	1613	2 G64488	reverse gyrase (in
68	60.7	1624	2 G71129	probable reverse g
69	60.7	2022	2 T48818	glucan 1,4-alpha-g
70	58.9	68	2 AG3181	hypothetical prote
71	58.9	146	2 F72458	hypothetical prote
72	58.9	179	2 D83266	molybdopterin bios
73	58.9	190	2 C75055	transcription regu
74	58.9	225	2 B83083	conserved hypothet
75	58.9	240	2 F71236	hypothetical prote
76	58.9	244	2 D75505	hypothetical prote
77	58.9	249	2 F91095	type III secretion
78	58.9	249	2 B85941	hypothetical prote
79	58.9	261	2 E82748	trypsinone reductas
80	58.9	272	2 A70955	hypothetical prote
81	58.9	275	2 C87634	conserved hypothet
82	58.9	287	2 H70923	probable hg transp
83	58.9	316	1 C69855	ABC transporter ho
84	58.9	357	2 E87584	transcription regu
85	58.9	360	2 C84243	asparagine synthet
86	58.9	384	2 AG3494	recF protein (impo
87	58.9	396	1 E64143	chloramphenicol re
88	58.9	396	2 E84196	cell division prot
89	58.9	432	2 B87161	competence damage
90	58.9	507	2 H86393	protein T24P13.8 l
91	58.9	547	2 T35143	probable monooxyge
92	58.9	547	2 G75364	DNA mismatch repai
93	58.9	630	2 T38023	probable transcrip
94	58.9	666	1 D69103	DNA helicase (EC 3
95	58.9	668	2 AG1989	hypothetical prote
96	58.9	679	2 S48939	hypothetical prote
97	58.9	707	2 A42322	ornithine decarbox
98	58.9	774	2 H83057	penicillin-binding
99	58.9	822	2 AB0238	cold-induced prote
100	58.9	822	2 AB0238	hemam storage sys
101	58.9	822	1 T47007	hypothetical prote
102	58.9	858	1 ZPECPA	penicillin-binding

103	33	58.9	858	2	A80999	penicillin-binding	176	32	57.1	1020	2	D83679	hypothetical prote
104	33	58.9	858	2	D86004	peptidoglycan synt	177	32	57.1	1036	2	A29832	HPI layer surface
105	33	58.9	858	2	F31158	peptidoglycan synt	178	32	57.1	1067	2	T18196	pol protein - silk
106	33	58.9	901	1	RGE1MT	regulatory protein	179	32	57.1	1146	2	B70376	reverse gyrase - A
107	33	58.9	901	2	D91161	positive regulator	180	32	57.1	1220	2	AD0125	exodeoxyribonuclea
108	33	58.9	901	2	D86007	positive regulator	181	32	57.1	1243	2	JC5615	membrane-associated
109	32.5	58.0	79	2	A87251	molybdopterin conv	182	32	57.1	1466	2	T30566	ATP-binding cassel
110	32.5	58.0	176	2	G70599	hypothetical prote	183	32	57.1	1621	2	S62356	TRP-195 protein -
111	32.5	58.0	398	2	A71482	hypothetical prote	184	32	57.1	1702	2	E86318	protein F15H18.4 [
112	32	57.1	73	2	F82172	hypothetical prote	185	32	57.1	2108	2	H70819	probable polyketid
113	32	57.1	75	2	S11156	psAS10 protein - c	186	32	57.1	3069	2	H70656	fatty-acid synthas
114	32	57.1	138	2	T49392	related to pullula	187	32	57.1	3076	2	A87058	fatty acid synthas
115	32	57.1	158	2	H83405	hypothetical prote	188	32	57.1	3104	2	S20473	fatty-acid synthas
116	32	57.1	176	2	T42179	hypothetical prote	189	32	57.1	3341	1	A42996	genome polyprotein
117	32	57.1	182	2	T24757	hypothetical prote	190	31.5	56.2	246	2	C81946	hypothetical prote
118	32	57.1	186	2	AF2205	hypothetical prote	191	31.5	56.2	250	2	C83397	pyrroloquinoline q
119	32	57.1	194	2	A75515	conserved hypotet	192	31.5	56.2	327	2	T35322	hypothetical prote
120	32	57.1	197	2	B71106	hypothetical prote	193	31.5	56.2	747	2	D95862	probable sensor hi
121	32	57.1	251	1	TRHUAZ	azurocidin precurs	194	31	55.4	57	2	C97937	transposase, uncha
122	32	57.1	261	2	D86029	probable CoA-trans	195	31	55.4	114	2	T30727	hypothetical prote
123	32	57.1	265	2	C83419	NADH-dependent eno	196	31	55.4	140	2	S20914	lysozyme (EC 3.2.1
124	32	57.1	286	2	H87506	hypothetical prote	197	31	55.4	140	2	S41577	lysozyme (EC 3.2.1
125	32	57.1	301	2	H84192	hypothetical prote	198	31	55.4	140	2	S41574	lysozyme (EC 3.2.1
126	32	57.1	306	2	T49541	hypothetical prote	199	31	55.4	141	2	S20915	lysozyme (EC 3.2.1
127	32	57.1	322	2	T11576	type IIIa membrane	200	31	55.4	148	1	FXDV	flavodoxin - Desul
128	32	57.1	324	2	H72637	hypothetical prote	201	31	55.4	151	2	F69931	hypothetical prote
129	32	57.1	335	2	A49654	transcription fact	202	31	55.4	157	2	B70892	hypothetical prote
130	32	57.1	346	2	T11577	type IIIa membrane	203	31	55.4	168	2	C70824	probable transcrip
131	32	57.1	360	2	T51394	reversibly glycosy	204	31	55.4	188	1	F64986	probable lipoprote
132	32	57.1	360	2	T33124	hypothetical prote	205	31	55.4	188	2	B85856	probable lipoprote
133	32	57.1	364	2	T04331	golgi associated p	206	31	55.4	188	2	C91012	probable lipoprote
134	32	57.1	364	2	T06507	reversibly glycosy	207	31	55.4	190	2	A70784	probable lipoprote
135	32	57.1	366	2	T72027	cell shape-determi	208	31	55.4	192	2	T28589	hypothetical prote
136	32	57.1	366	2	H86596	rod shape protein-	209	31	55.4	194	2	A36851	A27L protein - var
137	32	57.1	366	2	E81742	cell shape-determi	210	31	55.4	194	2	A72167	hypothetical prote
138	32	57.1	366	2	F71481	probable rod shape	211	31	55.4	200	2	B85047	hypothetical prote
139	32	57.1	366	2	AG0945	probable lipoprote	212	31	55.4	219	2	G84388	hypothetical prote
140	32	57.1	370	2	E71306	probable polynucle	213	31	55.4	221	2	D84516	hypothetical prote
141	32	57.1	370	2	C75268	carboxypeptidase G	214	31	55.4	226	2	F90469	hypothetical prote
142	32	57.1	370	2	B83480	probable oxidoredu	215	31	55.4	228	2	H95868	probable osmotical
143	32	57.1	371	2	T05213	hypothetical prote	216	31	55.4	229	2	H75594	hypothetical prote
144	32	57.1	378	2	T02895	hypothetical prote	217	31	55.4	240	2	D75203	hypothetical prote
145	32	57.1	383	2	F83321	hypothetical prote	218	31	55.4	241	2	B84992	peptidylprolyl iso
146	32	57.1	392	2	AD2360	hypothetical prote	219	31	55.4	247	2	A96001	conserved hypotet
147	32	57.1	402	2	AC2039	hypothetical prote	220	31	55.4	253	2	G86605	snGlycerol-3-P acy
148	32	57.1	434	2	F84332	succinoglycan bios	221	31	55.4	253	2	H72017	snGlycerol-3-p acy
149	32	57.1	447	2	F69433	conserved hypotet	222	31	55.4	255	2	AB3497	methyitransferase
150	32	57.1	465	2	A83536	penicillin amidase	223	31	55.4	256	2	B75099	hypothetical prote
151	32	57.1	477	2	A87285	phosphate ABC tran	224	31	55.4	257	2	G70788	hypothetical prote
152	32	57.1	480	2	T50511	serine-type carbox	225	31	55.4	280	2	H70942	hypothetical prote
153	32	57.1	509	2	S74935	hypothetical prote	226	31	55.4	280	2	T03543	probable cobalt tr
154	32	57.1	509	2	S70825	AFG1 protein - yea	227	31	55.4	285	2	S65838	dihydroperoxide sy
155	32	57.1	539	2	B95313	cytochrome-c oxida	228	31	55.4	285	2	C84170	ribosomal protein
156	32	57.1	544	2	T05211	hypothetical prote	229	31	55.4	292	2	A75409	carboxymethylenebu
157	32	57.1	574	2	T05214	hypothetical prote	230	31	55.4	299	2	B87049	hypothetical prote
158	32	57.1	589	2	A82499	hypothetical prote	231	31	55.4	301	1	VHBPLL	major capsid prote
159	32	57.1	592	2	T48346	photoreceptor-inte	232	31	55.4	301	2	T03320	hypothetical prote
160	32	57.1	649	2	F64847	epithelial sodium	233	31	55.4	302	2	S55644	hypothetical prote
161	32	57.1	658	2	H95852	hypothetical fusio	234	31	55.4	310	2	C83267	hypothetical prote
162	32	57.1	669	2	F84433	hypothetical prote	235	31	55.4	318	2	G84143	hypothetical prote
163	32	57.1	690	2	C89208	DNA helicase relat	236	31	55.4	330	2	E98119	transposase, uncha
164	32	57.1	699	2	T09069	probable cAMP-resp	237	31	55.4	343	2	D75260	probable Fe-s-clus
165	32	57.1	709	2	T26874	hypothetical prote	238	31	55.4	345	2	A87113	xenobiotic reducta
166	32	57.1	727	2	F95472	catalase (EC 1.1.1	239	31	55.4	350	2	D83102	conserved hypotet
167	32	57.1	741	2	E75428	hypothetical prote	240	31	55.4	363	2	AC0244	5S RNA-binding pro
168	32	57.1	749	2	B75422	ferrous iron trans	241	31	55.4	365	2	A34895	probable ABC trans
169	32	57.1	818	2	T12975	outer envelope mem	242	31	55.4	372	1	G70817	protidine-5'-phosp
170	32	57.1	840	2	F84213	hypothetical prote	243	31	55.4	376	1	DCCEOC	hypothetical prote
171	32	57.1	870	2	B82732	glycerol-3-phospha	244	31	55.4	386	2	G70779	hypothetical prote
172	32	57.1	892	2	A87447	sensor histidine k	245	31	55.4	399	2	T40877	hypothetical prote
173	32	57.1	935	2	E96806	hypothetical prote	246	31	55.4	400	2	T46526	probable hydroxyla
174	32	57.1	994	2	A48849	Ca2+-transporting	247	31	55.4	410	2	B55523	aromatic dioxygena
175	32	57.1	1001	1	PWRBFC	Ca2+-transporting	248	31	55.4	412	2	C86177	hypothetical prote

249	31	55.4	415	2	D96759	probable serine ca	322	31	55.4	1343	2	AF0611	cell division prot
250	31	55.4	416	2	B64132	tRNA adenylyltrans	323	31	55.4	1516	2	F83085	conserved hypotet
251	31	55.4	431	2	AG1884	glycolate oxidase	324	31	55.4	1564	2	S55517	probable transport
252	31	55.4	432	2	T36482	probable aminopept	325	31	55.4	1582	2	E70876	probable polyketid
253	31	55.4	433	2	B95405	probable nitrilotri	326	31	55.4	1619	2	T30541	ABC1 transport pro
254	31	55.4	435	2	E84618	probable serine ca	327	31	55.4	1647	2	T49412	hypothetical prote
255	31	55.4	440	2	B87311	conserved hypotet	328	31	55.4	1974	2	T30010	hypothetical prote
256	31	55.4	441	2	A96759	protein serine car	329	31	55.4	4063	2	T42993	probable spectrin
257	31	55.4	443	2	S66040	serine-type D-Ala-	330	31	55.4	4101	2	T23630	hypothetical prote
258	31	55.4	443	2	JC4088	chorionic gonadotr	331	31	55.4	15281	2	S41309	cyclosporin synthe
259	31	55.4	443	2	T45574	anthranilate N-hyd	332	30.5	54.5	333	1	A38094	D-lactate dehydrog
260	31	55.4	448	2	A55519	mcrA protein - Str	333	30.5	54.5	337	1	S29296	probable substrate
261	31	55.4	468	2	H70521	probable papA2 pro	334	30.5	54.5	542	2	T35938	55K protein precu
262	31	55.4	468	2	DYH205	dopamine receptor	335	30.5	54.5	542	2	S07386	hypothetical prote
263	31	55.4	477	1	WZBE20	gene 20 protein -	336	30.5	54.5	583	2	H69165	hypothetical prote
264	31	55.4	492	2	B86253	hypothetical prote	337	30.5	54.5	604	2	T36966	hypothetical prote
265	31	55.4	496	2	B84974	N utilization subs	338	30.5	54.5	611	2	F82442	hypothetical prote
266	31	55.4	527	2	A70975	hypothetical prote	339	30.5	54.5	639	2	H70661	probable peptide A
267	31	55.4	532	2	A35149	ipah protein - Shi	340	30.5	54.5	1171	2	F81142	probable dnag prot
268	31	55.4	543	2	T32109	hypothetical prote	341	30	53.6	46	2	F81142	exodeoxyribonucle
269	31	55.4	554	2	E83071	probable ATP-bind	342	30	53.6	46	2	F81142	hypothetical prote
270	31	55.4	555	2	G86138	ABC transporter in	343	30	53.6	50	2	S40969	TyB protein - yeas
271	31	55.4	555	2	E91297	ABC transporter in	344	30	53.6	80	2	A60699	hypothetical prote
272	31	55.4	555	2	F65254	ABC transporter in	345	30	53.6	101	2	D69745	hypothetical prote
273	31	55.4	555	2	AE1074	conserved hypotet	346	30	53.6	104	2	B84037	hypothetical prote
274	31	55.4	556	2	G64169	potassium-depend	347	30	53.6	105	2	A97519	hypothetical prote
275	31	55.4	561	2	AG2336	acyl-CoA dehydrog	348	30	53.6	120	2	D83166	hypothetical prote
276	31	55.4	564	2	E87631	RNA binding protei	349	30	53.6	127	2	S40034	killer toxin KP4 p
277	31	55.4	565	2	T38941	ipah protein - Shi	350	30	53.6	130	2	AC2735	hypothetical prote
278	31	55.4	574	2	B35149	gamma-glutamyltran	351	30	53.6	136	2	H84358	hypothetical prote
279	31	55.4	580	2	AB0994	archaeosine tRNA-r	352	30	53.6	150	2	AC3217	conserved hypotet
280	31	55.4	585	2	F84346	2,3-dihydroxybenzo	353	30	53.6	158	2	F95415	hypothetical prote
281	31	55.4	587	2	AD3519	hypothetical prote	354	30	53.6	160	2	AB2020	probable periplasm
282	31	55.4	594	2	T31744	probable ABC-type	355	30	53.6	160	2	D81247	thioredoxin-relate
283	31	55.4	601	2	T36411	hemagglutinin - ph	356	30	53.6	162	2	D97821	hypothetical prote
284	31	55.4	607	1	HM22PD	hemagglutinin - ph	357	30	53.6	184	2	G87321	hypothetical prote
285	31	55.4	607	1	JQ1535	chaperone protein	358	30	53.6	189	2	A85489	hypothetical prote
286	31	55.4	616	2	AB0824	probable oxidoredu	359	30	53.6	189	2	A99638	hypothetical prote
287	31	55.4	648	2	C83023	TonB-dependent rec	360	30	53.6	194	2	AI0155	probable lipoprote
288	31	55.4	660	2	D87331	hypothetical prote	361	30	53.6	209	2	AF2037	hypothetical prote
289	31	55.4	690	2	T47478	hypothetical prote	362	30	53.6	219	1	TRFGAZ	azurocidin - pig
290	31	55.4	698	2	T51915	hypothetical prote	363	30	53.6	227	2	B89057	protein K09H11.2 [
291	31	55.4	720	2	S71130	catalase (EC 1.11.	364	30	53.6	231	2	G70000	hypothetical prote
292	31	55.4	722	2	D89701	protein F28H6.4 [i	365	30	53.6	234	2	T26560	hypothetical prote
293	31	55.4	733	2	G84668	hypothetical prote	366	30	53.6	235	2	AF2319	glycerol uptake fa
294	31	55.4	735	2	F85055	hypothetical prote	367	30	53.6	238	2	C86656	lmby protein - Str
295	31	55.4	743	2	T32794	hypothetical prote	368	30	53.6	246	2	AI1717	hypothetical prote
296	31	55.4	754	2	S75113	catalase (EC 1.11.	369	30	53.6	249	2	F83477	hypothetical prote
297	31	55.4	771	2	A83348	probable aldehyde	370	30	53.6	255	2	T21686	conserved hypotet
298	31	55.4	781	2	F82584	catalase/peroxidas	371	30	53.6	257	2	D87516	conserved hypotet
299	31	55.4	836	2	S55344	outer envelope mem	372	30	53.6	262	1	A31372	probable enoyl-CoA
300	31	55.4	836	2	S49940	cell division cont	373	30	53.6	263	2	D83289	beta-lactamase (EC
301	31	55.4	885	1	S78246	endopeptidase Clp	374	30	53.6	264	2	I39696	probable DNA-(apur
302	31	55.4	903	2	AI0015	maltose regulon po	375	30	53.6	264	2	C72770	conserved hypotet
303	31	55.4	907	2	S54353	inter-alpha-trypsi	376	30	53.6	266	2	AI2762	hypothetical prote
304	31	55.4	909	2	T06635	hypothetical prote	377	30	53.6	267	2	E70837	hypothetical prote
305	31	55.4	911	2	A39967	inter-alpha-trypsi	378	30	53.6	271	2	H97543	MEP4 protein - yea
306	31	55.4	916	2	T35680	probable regulatory	379	30	53.6	273	2	S48511	hypothetical prote
307	31	55.4	921	2	D82513	malt regulatory pr	380	30	53.6	274	2	D97653	conserved hypotet
308	31	55.4	944	2	C70839	probable mmpL3 pro	381	30	53.6	274	2	AB2877	magz protein [impo
309	31	55.4	947	2	H85088	hypothetical prote	382	30	53.6	275	2	G87465	oxidoreductase of
310	31	55.4	983	2	G84524	probable disease r	383	30	53.6	276	2	T45275	probable Beta lact
311	31	55.4	993	2	F97717	hypothetical prote	384	30	53.6	281	2	D95395	succinoglycan bios
312	31	55.4	1000	2	T21522	hypothetical prote	385	30	53.6	281	2	E49348	hypothetical prote
313	31	55.4	1038	1	WZBE21	gene 21 protein -	386	30	53.6	283	2	S73935	transcription regu
314	31	55.4	1042	2	S43904	hyaluronidase - Cl	387	30	53.6	289	2	F69897	conserved hypotet
315	31	55.4	1080	2	E87586	metal ion efflux R	388	30	53.6	290	2	F64558	hypothetical prote
316	31	55.4	1120	2	H88449	protein F54D8.1 [i	389	30	53.6	293	2	H71949	hypothetical prote
317	31	55.4	1140	2	T09486	hypothetical prote	390	30	53.6	293	2	F69897	conserved hypotet
318	31	55.4	1178	2	AC3394	DNA-directed DNA p	391	30	53.6	302	2	AH0191	probable polysacch
319	31	55.4	1329	2	A64858	cell division prot	392	30	53.6	303	2	T13598	trypsin homolog -
320	31	55.4	1342	2	E85614	cell division prot	393	30	53.6	303	2	B47089	probable Arac-type
321	31	55.4	1342	2	G90750	cell division prot	394	30	53.6	308	2	E98259	probable transcrip

395	30	53.6	308	2	AD3025	transcription regu	468	30	53.6	427	2	H87571	N-carbamyl-L-amino
396	30	53.6	312	2	T06473	probable obfusol	469	30	53.6	428	2	T03251	calnexin - maize (
397	30	53.6	315	2	T45928	reductase-like pro	470	30	53.6	433	2	A84619	probable serine ca
398	30	53.6	316	2	G82557	hypothetical prote	471	30	53.6	433	2	B90288	thiamin biosynthes
399	30	53.6	319	2	T01822	hypothetical prote	472	30	53.6	433	2	H90359	thiamin biosynthes
400	30	53.6	323	2	D72661	hypothetical prote	473	30	53.6	435	1	T07310	prochlorophyllid
401	30	53.6	328	2	G95392	probable protease	474	30	53.6	436	2	B86469	protein F12K21.16
402	30	53.6	334	2	A83662	transcription regu	475	30	53.6	437	2	G84505	probable PttA2-lik
403	30	53.6	334	2	D86455	T9L6.3 protein - A	476	30	53.6	439	2	T28841	hypothetical prote
404	30	53.6	334	2	AE0154	probable nucleoid-	477	30	53.6	444	2	D86283	T15D22.4 protein -
405	30	53.6	335	2	AC0786	nucleoid-associated	478	30	53.6	445	2	A49189	cholesterol monoox
406	30	53.6	335	2	A64988	hypothetical 37.8	479	30	53.6	446	2	F84506	En/Spm-like transp
407	30	53.6	335	2	H85857	nucleoid-associate	480	30	53.6	446	2	C75087	hydroxymethylpyrim
408	30	53.6	335	2	F31013	nucleoid-associate	481	30	53.6	453	2	T06475	probable obfusol
409	30	53.6	338	2	S67158	26S proteasome reg	482	30	53.6	457	2	T51660	dopamine D1B recep
410	30	53.6	338	2	H83228	hypothetical prote	483	30	53.6	459	2	G84752	En/Spm-like transp
411	30	53.6	343	2	F72319	translation initia	484	30	53.6	466	1	YTB5MA	methylethionine A r
412	30	53.6	345	2	AG0366	thiosulfate-bindin	485	30	53.6	468	2	D84503	probable serine ca
413	30	53.6	347	2	AD2711	aldo/keto reductas	486	30	53.6	473	2	T05306	homeobox protein A
414	30	53.6	355	2	AE0518	hypothetical prote	487	30	53.6	476	2	C85041	probable DNA-bindi
415	30	53.6	356	2	E84495	En/Spm-like transp	488	30	53.6	488	2	G81295	cytochrome-c oxida
416	30	53.6	360	2	A86175	hypothetical prote	489	30	53.6	502	2	A45737	phenylcarbamate hy
417	30	53.6	362	2	AC3192	hypothetical prote	490	30	53.6	502	2	AF3222	conserved hypothet
418	30	53.6	367	2	F86315	hypothetical prote	491	30	53.6	517	2	E95959	probable membrane-
419	30	53.6	370	2	B97493	hypothetical prote	492	30	53.6	520	1	O4BOM	cholesterol monoox
420	30	53.6	371	2	A86264	protein F3P19.3 li	493	30	53.6	520	1	S03188	cholesterol monoox
421	30	53.6	372	2	T14193	L-ascorbate peroxi	494	30	53.6	531	2	C95338	hypothetical prote
422	30	53.6	372	2	G64064	outer membrane int	495	30	53.6	538	2	A70836	hypothetical prote
423	30	53.6	373	2	T47115	probable 4-carboxy	496	30	53.6	538	2	T27433	hypothetical prote
424	30	53.6	376	2	T34734	hypothetical prote	497	30	53.6	538	2	H83466	probable ATP-bindi
425	30	53.6	376	2	A02585	DNA replication an	498	30	53.6	541	2	B87532	transporter, proba
426	30	53.6	376	2	C97367	recF protein limpo	499	30	53.6	544	2	B83347	hypothetical prote
427	30	53.6	385	2	G87519	hypothetical prote	500	30	53.6	550	2	B70583	probable transposa
428	30	53.6	387	2	S32934	aminotransferase p	501	30	53.6	552	2	E87226	conserved membrane
429	30	53.6	389	2	T47323	hypothetical prote	502	30	53.6	570	2	D96766	protein glucosyltr
430	30	53.6	392	1	A24608	gastriclin (FC 3.4	503	30	53.6	575	1	S52819	isocitrate lyase (
431	30	53.6	396	2	AE0677	probable membrane	504	30	53.6	577	2	E90044	hypothetical prote
432	30	53.6	398	2	H87485	lipid-A-disacchari	505	30	53.6	578	2	T33916	hypothetical prote
433	30	53.6	398	2	B70415	proton/sodium-glut	506	30	53.6	578	2	AC0119	probable exported
434	30	53.6	402	2	A10323	transposase, IS285	507	30	53.6	578	2	AC0179	probable ompa-fam
435	30	53.6	402	2	AB0193	transposase, IS285	508	30	53.6	590	2	T45820	hypothetical prote
436	30	53.6	402	2	T14710	probable transposa	509	30	53.6	591	1	S30145	ketol-acid reducto
437	30	53.6	402	2	AE0267	transposase, IS285	510	30	53.6	591	1	T45681	ketol-acid reducto
438	30	53.6	402	2	A10256	transposase, IS285	511	30	53.6	594	2	I48771	Slp(w7) - mouse (f
439	30	53.6	402	2	AH0298	transposase, IS285	512	30	53.6	595	1	S17160	ketol-acid reducto
440	30	53.6	402	2	AG0264	transposase, IS285	513	30	53.6	596	2	C44611	voltage-dependent
441	30	53.6	402	2	AB0472	transposase, IS285	514	30	53.6	596	2	G75457	tetratricopeptide
442	30	53.6	402	2	AB0197	transposase, IS285	515	30	53.6	597	2	I52859	L-type voltage-gat
443	30	53.6	402	2	AE0190	transposase, IS285	516	30	53.6	597	2	S18304	calcium channel pr
444	30	53.6	402	2	AG0210	transposase, IS285	517	30	53.6	604	1	HMNZCD	hemagglutinin - ca
445	30	53.6	402	2	AG0339	transposase, IS285	518	30	53.6	607	2	S49017	hemagglutinin - ca
446	30	53.6	402	2	AB0358	transposase, IS285	519	30	53.6	615	1	KFHU12	cosagulation factor
447	30	53.6	402	2	AH0181	transposase, (impor	520	30	53.6	629	2	F86351	protein T26F17.2 [
448	30	53.6	402	2	AE0193	transposase, IS285	521	30	53.6	651	2	T10219	protein kinase hom
449	30	53.6	402	2	AE0242	transposase, IS285	522	30	53.6	653	2	D95848	probable elongatio
450	30	53.6	402	2	AG0260	transposase, IS285	523	30	53.6	664	2	F83376	conserved hypothet
451	30	53.6	402	2	AC0341	transposase, IS285	524	30	53.6	666	2	T35864	hypothetical prote
452	30	53.6	402	2	AD0002	transposase, IS285	525	30	53.6	671	2	H91045	hypothetical prote
453	30	53.6	402	2	AD0313	transposase, IS285	526	30	53.6	671	2	D85890	hypothetical prote
454	30	53.6	402	2	AF0348	transposase, IS285	527	30	53.6	671	2	A65023	hypothetical prote
455	30	53.6	406	2	T47674	hypothetical prote	528	30	53.6	674	2	F82844	oligopeptidase A X
456	30	53.6	407	2	G82040	general secretion	529	30	53.6	680	2	H84774	probable homeodoma
457	30	53.6	408	2	B87436	conserved hypothet	530	30	53.6	682	1	S22700	amphipysin - chic
458	30	53.6	408	2	C96502	hypothetical prote	531	30	53.6	683	2	A83674	hypothetical prote
459	30	53.6	410	2	B84176	oxalate/formate an	532	30	53.6	689	2	B69135	coenzyme F420-redu
460	30	53.6	412	2	C72391	conserved hypothet	533	30	53.6	694	2	T01504	hypothetical prote
461	30	53.6	414	2	S43253	alanine-glyoxylate	534	30	53.6	711	2	T12525	hypothetical prote
462	30	53.6	421	2	T10190	L-ascorbate peroxi	535	30	53.6	712	2	E83067	probable chemotaxi
463	30	53.6	421	2	S71331	transposase, IS285	536	30	53.6	724	2	C49423	semaphorin II prec
464	30	53.6	423	2	T01319	hypothetical prote	537	30	53.6	773	2	T39597	hypothetical prote
465	30	53.6	423	2	S48817	hypothetical prote	538	30	53.6	777	2	T44597	transducer protein
466	30	53.6	425	2	F70608	hypothetical prote	539	30	53.6	778	2	F84237	Htr4 transducer [i
467	30	53.6	426	2	C96804	hypothetical prote	540	30	53.6	778	2	T48897	transducer protein

541	30	53.6	791	2	AC0017	conserved hypothet	614	30	53.6	1755	2	S69839	TyB protein - yeas
542	30	53.6	822	2	C97211	uncharacterized co	615	30	53.6	1755	2	S69845	TyB protein - yeas
543	30	53.6	861	2	T36381	probable large ATP	616	30	53.6	1755	2	S69980	TyB protein - yeas
544	30	53.6	875	2	T27495	hypothetical prote	617	30	53.6	1756	2	S45867	TyB protein - yeas
545	30	53.6	882	2	AF3036	nitrate reductase	618	30	53.6	1756	2	S69983	TyB protein - yeas
546	30	53.6	892	2	F87325	hypothetical prote	619	30	53.6	1770	2	S69953	TyB protein - yeas
547	30	53.6	899	2	T35299	probable pyruvate	620	30	53.6	1770	2	S58651	TyB protein - yeas
548	30	53.6	903	2	E98249	nitrate reductase	621	30	53.6	1770	2	S70233	TyB protein - yeas
549	30	53.6	906	1	RNBPX1	DNA-directed RNA p	622	30	53.6	1770	2	S69948	TyB protein - yeas
550	30	53.6	911	2	T08105	nitrate reductase	623	30	53.6	1770	2	S69966	TyB protein - yeas
551	30	53.6	911	2	T08108	nitrate reductase	624	30	53.6	1770	2	S45842	TyB protein - yeas
552	30	53.6	917	2	E96807	nitrate reductase	625	30	53.6	1770	2	S70230	TyB protein - yeas
553	30	53.6	917	2	S35228	nitrate reductase	626	30	53.6	1770	2	S69950	TyB protein - yeas
554	30	53.6	919	2	S37786	hypothetical prote	627	30	53.6	1793	2	S52601	TyB protein - yeas
555	30	53.6	976	2	C87389	TonB-dependent rec	628	30	53.6	1810	1	S69973	TyB protein - yeas
556	30	53.6	1026	1	TLBP71	tail fiber protein	629	30	53.6	1968	1	PNQ093	genome polyprotein
557	30	53.6	1054	2	H69377	reverse gyrase (co	630	30	53.6	1997	1	S12050	protein-tyrosine-p
558	30	53.6	1063	2	T18255	cytoskeleton assem	631	30	53.6	2163	2	T51397	hypothetical prote
559	30	53.6	1063	2	T18255	probable isoleucyl	632	30	53.6	2264	1	GNVVTB	genome polyprotein
560	30	53.6	1064	2	E72729	dyein heavy chain	633	30	53.6	2948	2	T26664	hypothetical prote
561	30	53.6	1078	2	T30879	RNA-directed RNA p	634	30	53.6	3450	2	T26963	hypothetical prote
562	30	53.6	1082	2	T13410	hypothetical prote	635	30	53.6	3461	2	T36964	hypothetical prote
563	30	53.6	1124	2	F71719	hypothetical prote	636	30	53.6	4151	2	G70944	probable polyketid
564	30	53.6	1134	2	T04587	hypothetical prote	637	30	53.6	4464	2	D87755	protein T2E12.4 [
565	30	53.6	1164	2	T06144	disease resistance	638	30	53.6	4464	2	D87755	protein T2E12.4 [
566	30	53.6	1196	2	S40908	TyB protein - yeas	639	29.5	52.7	121	2	T42138	type II secretion
567	30	53.6	1198	2	D96723	hypothetical prote	640	29.5	52.7	141	2	AF3317	hypothetical prote
568	30	53.6	1209	2	E90994	probable regulator	641	29.5	52.7	161	2	T35350	hypothetical prote
569	30	53.6	1209	2	H85839	probable regulator	642	29.5	52.7	251	2	G75440	conserved hypothet
570	30	53.6	1210	2	E64979	hypothetical 138.1	643	29.5	52.7	315	2	S58171	bifunctional cycla
571	30	53.6	1212	2	F83153	probable two-compo	644	29.5	52.7	341	2	E97181	guanine synthetase
572	30	53.6	1243	2	S07278	tail fiber protein	645	29.5	52.7	356	2	A11954	hypothetical prote
573	30	53.6	1278	2	B22671	TyB protein - yeas	646	29.5	52.7	383	2	T48870	muconate cycloisom
574	30	53.6	1282	2	T30577	DNA topoisomerase	647	29.5	52.7	385	2	T46824	muconate cycloisom
575	30	53.6	1328	2	B22999	TyB protein - yeas	648	29.5	52.7	646	2	A70903	probable exinucle
576	30	53.6	1328	2	S52481	TyB protein - yeas	649	29.5	52.7	647	2	B86979	ABC excision nucle
577	30	53.6	1328	2	S50948	TyB protein - yeas	650	29.5	52.7	728	2	T36026	NADP-dependent ald
578	30	53.6	1328	2	S52894	TyB protein - yeas	651	29.5	52.7	788	2	AE3153	hypothetical prote
579	30	53.6	1328	2	B28097	TyB protein - yeas	652	29.5	52.7	798	2	F98134	genome polyprotein
580	30	53.6	1348	2	B23496	TyB protein - yeas	653	29.5	52.7	3433	1	GNWVKV	probable non-ribos
581	30	53.6	1375	2	T13822	frazzled gene prot	654	29.5	52.7	5149	2	F83345	hypothetical prote
582	30	53.6	1376	2	F72609	probable reverse g	655	29	51.8	50	2	H97798	rubredoxin PAS351
583	30	53.6	1526	2	T13823	frazzled gene prot	656	29	51.8	55	2	AE2977	hypothetical prote
584	30	53.6	1538	2	B90524	probable ATP-depen	657	29	51.8	67	2	S20172	hypothetical prote
585	30	53.6	1538	2	F85772	ATP-dependent heli	658	29	51.8	73	2	G82382	probable ribosomal
586	30	53.6	1538	2	G64922	probable phosphati	659	29	51.8	74	2	B72512	hypothetical prote
587	30	53.6	1593	2	T14176	probable ATP-depen	660	29	51.8	74	2	C81133	secretory protein
588	30	53.6	1598	2	S69967	TyB protein - yeas	661	29	51.8	81	2	S77318	hypothetical prote
589	30	53.6	1600	2	AB3281	glutamate dehydrog	662	29	51.8	93	2	D82474	conserved hypothet
590	30	53.6	1690	2	T31670	DNA-directed RNA p	663	29	51.8	102	2	C96034	hypothetical prote
591	30	53.6	1735	2	S54784	sex-limited protei	664	29	51.8	110	2	D90328	30S ribosomal prot
592	30	53.6	1736	2	A29176	sex-limited protei	665	29	51.8	121	2	E90019	hypothetical prote
593	30	53.6	1738	1	A24558	complement C4 prec	666	29	51.8	126	2	S35585	hypothetical prote
594	30	53.6	1749	2	S69972	TyB protein - yeas	667	29	51.8	128	2	S72600	hypothetical prote
595	30	53.6	1755	2	S69979	TyB protein - yeas	668	29	51.8	130	2	C40368	lg kappa chain - h
596	30	53.6	1755	2	S69975	TyB protein - yeas	669	29	51.8	131	2	C72680	hypothetical prote
597	30	53.6	1755	2	S69866	TyB protein - yeas	670	29	51.8	133	2	H75386	hypothetical prote
598	30	53.6	1755	2	S69955	TyB protein - yeas	671	29	51.8	137	2	AB3157	conserved hypothet
599	30	53.6	1755	2	S50663	TyB protein - yeas	672	29	51.8	137	2	G98130	probable gntR-fami
600	30	53.6	1755	2	S50641	TyB protein - yeas	673	29	51.8	146	2	T36436	probable PTS syste
601	30	53.6	1755	2	S57047	TyB protein - yeas	674	29	51.8	148	2	F71367	probable capsid co
602	30	53.6	1755	2	S69951	TyB protein - yeas	675	29	51.8	154	2	AD1037	hypothetical prote
603	30	53.6	1755	2	S45736	TyB protein - yeas	676	29	51.8	158	2	AD1902	hypothetical prote
604	30	53.6	1755	2	S69838	TyB protein - yeas	677	29	51.8	164	2	A87556	hypothetical prote
605	30	53.6	1755	2	S61763	TyB protein - yeas	678	29	51.8	165	2	C97979	hypothetical prote
606	30	53.6	1755	2	S57045	TyB protein - yeas	679	29	51.8	178	2	F90168	hypothetical prote
607	30	53.6	1755	2	S70238	TyB protein - yeas	680	29	51.8	179	2	S6835	unknown protein F5
608	30	53.6	1755	2	S69982	TyB protein - yeas	681	29	51.8	179	2	S39078	auxin-induced prot
609	30	53.6	1755	2	S69960	TyB protein - yeas	682	29	51.8	181	1	T1SYB	trypsin inhibitor
610	30	53.6	1755	2	S69969	TyB protein - yeas	683	29	51.8	185	2	T10677	pathogenesis-relat
611	30	53.6	1755	2	S69949	TyB protein - yeas	684	29	51.8	185	2	A70682	hypothetical prote
612	30	53.6	1755	2	S69957	TyB protein - yeas	685	29	51.8	188	2	T10859	auxin-induced prot
613	30	53.6	1755	2	S69963	TyB protein - yeas	686	29	51.8	193	2	H95959	hypothetical membr

687	29	51.8	194	2	T10939	auxin-induced prot	760	29	51.8	312	2	S59519	tryptophan synthas
688	29	51.8	195	2	E95110	ABC transporter, A	761	29	51.8	319	2	S62196	hypothetical prote
689	29	51.8	199	2	H87416	phosphoglycerate m	762	29	51.8	324	2	D95869	probable nutrient
690	29	51.8	200	2	T28839	hypothetical prote	763	29	51.8	326	2	T51675	ribonucleoprotein
691	29	51.8	201	2	T08333	conserved hypot het	764	29	51.8	327	2	T21873	hypothetical prote
692	29	51.8	203	2	J01091	trypsin inhibitor	765	29	51.8	330	2	A86921	antigen 85A, mycol
693	29	51.8	204	2	T41513	probable proteasom	766	29	51.8	330	2	D95362	probable Arac-fami
694	29	51.8	204	2	J01092	trypsin inhibitor	767	29	51.8	331	2	S76480	hypothetical prote
695	29	51.8	204	2	T37226	hypothetical prote	768	29	51.8	331	2	MMBEK1	cell fusion protei
696	29	51.8	205	2	S29251	hypothetical prote	769	29	51.8	338	1	MMBEK2	cell fusion protei
697	29	51.8	206	2	AB2759	Conserved hypot het	770	29	51.8	338	1	S10326	alpha-antigen A, e
698	29	51.8	206	2	AB2546	hypothetical prote	771	29	51.8	338	2	H70887	32k antigen fbpa p
699	29	51.8	209	2	S54133	hypothetical 23.7K	772	29	51.8	339	2	S76619	hypothetical prote
700	29	51.8	211	2	T17228	hypothetical prote	773	29	51.8	341	2	R86483	unknown protein, 4
701	29	51.8	214	2	B40256	hypothetical prote	774	29	51.8	341	2	R87620	conserved hypot het
702	29	51.8	215	2	B83429	conserved hypot het	775	29	51.8	346	2	D64484	hypothetical prote
703	29	51.8	216	2	AF0160	probable hydrolase	776	29	51.8	347	2	H82391	NADH-dependent fla
704	29	51.8	217	2	S19190	trypsin inhibitor	777	29	51.8	349	2	H72715	probable aspartate
705	29	51.8	218	2	T35174	hypothetical prote	778	29	51.8	350	2	T50581	binding-protein-de
706	29	51.8	221	2	D86600	SET domain protein	779	29	51.8	351	2	AH2205	hypothetical prote
707	29	51.8	221	2	B81514	conserved hypot het	780	29	51.8	352	2	D90390	conserved hypot het
708	29	51.8	223	2	D72024	set domain protein	781	29	51.8	362	2	T17477	probable 2-dehydro
709	29	51.8	223	2	S77461	hypothetical prote	782	29	51.8	365	2	C86647	hypothetical prote
710	29	51.8	226	2	D83249	probable hydrolase	783	29	51.8	366	2	A99375	transport protein,
711	29	51.8	226	2	H30233	dtpp-4-dehydroxam	784	29	51.8	366	2	S74005	conserved hypot het
712	29	51.8	227	2	T45388	hypothetical prote	785	29	51.8	370	2	S75625	hypothetical prote
713	29	51.8	229	2	F84370	hypothetical prote	786	29	51.8	370	2	C70747	hypothetical prote
714	29	51.8	230	2	D75518	probable biopolyme	787	29	51.8	371	2	H87374	perosamine synthet
715	29	51.8	236	2	E84255	hypothetical prote	788	29	51.8	371	2	A13558	glycerol trinitrat
716	29	51.8	238	2	G87267	ribonuclease PH [i	789	29	51.8	372	2	A12862	oxidoreductase Atu
717	29	51.8	238	2	E70524	hypothetical prote	790	29	51.8	375	2	S76663	hypothetical prote
718	29	51.8	239	2	B83296	probable transcrip	791	29	51.8	376	2	G83290	hypothetical prote
719	29	51.8	242	2	A22681	conserved hypot het	792	29	51.8	379	2	A80959	hypothetical prote
720	29	51.8	246	2	H97539	hypothetical prote	793	29	51.8	381	2	D75569	2-oxo acid dehydro
721	29	51.8	251	2	S23821	hypothetical prote	794	29	51.8	381	2	F70669	coenzyme F420-depe
722	29	51.8	251	2	C45557	regulatory protein	795	29	51.8	382	2	C86925	probable oxidoredu
723	29	51.8	251	2	F83344	probable ATP-bindi	796	29	51.8	383	2	C95965	probable two-compo
724	29	51.8	253	2	T41139	proteasome compone	797	29	51.8	385	2	S44232	strH protein - Str
725	29	51.8	257	1	B64042	formate dehydrogen	798	29	51.8	386	2	H75287	hypothetical prote
726	29	51.8	257	1	S6108	26S proteasome reg	799	29	51.8	386	2	AE1996	hypothetical prote
727	29	51.8	257	2	B86987	conserved hypot het	800	29	51.8	386	2	T36425	probable sugar iso
728	29	51.8	258	2	B36914	probable nicotinac	801	29	51.8	388	2	H72223	thiamin biosynthes
729	29	51.8	258	2	D95910	probable membrane-	802	29	51.8	389	2	E81806	O-succinylhomoseri
730	29	51.8	259	2	A87623	PAP2 homolog prote	803	29	51.8	390	2	H87345	thiolase family pr
731	29	51.8	259	2	A96740	Fl4023.14 [importe	804	29	51.8	390	2	AE3384	N-acetylmuramoyl-L
732	29	51.8	260	2	T01837	hypothetical prote	805	29	51.8	392	2	G85076	probable reverse t
733	29	51.8	263	1	ASLJFP	vif protein - feli	806	29	51.8	393	1	C70929	cytochrome P450 Kv
734	29	51.8	264	2	F96807	unknown protein T3	807	29	51.8	394	2	JC5197	aromatic-amino-aci
735	29	51.8	264	2	JU0328	cellulase (EC 3.2.	808	29	51.8	394	2	F72517	hypothetical prote
736	29	51.8	265	2	C70221	conserved hypot het	809	29	51.8	396	1	C64907	chloramphenicol re
737	29	51.8	267	2	H96668	protein FIN19.9 [i	810	29	51.8	396	2	F83130	probable MFS trans
738	29	51.8	272	1	C64009	ribonuclease homol	811	29	51.8	396	2	H85721	probable resistanc
739	29	51.8	274	2	C85084	hypothetical prote	812	29	51.8	396	2	G90895	hypothetical prote
740	29	51.8	275	2	C70924	conserved hypot het	813	29	51.8	399	2	B82197	zona occludens tox
741	29	51.8	276	2	T08178	assemblin, striate	814	29	51.8	399	2	A43864	zonula occludens t
742	29	51.8	280	2	B59931	probable amino aci	815	29	51.8	400	2	B70806	probable acyl-CoA
743	29	51.8	283	2	T00581	hypothetical prote	816	29	51.8	401	2	B82868	plasmid replicatio
744	29	51.8	285	2	A87350	conserved hypot het	817	29	51.8	403	2	S76603	hypothetical prote
745	29	51.8	290	2	I64051	phosphatidylserine	818	29	51.8	409	2	G86257	hypothetical prote
746	29	51.8	291	2	G82136	hypothetical prote	819	29	51.8	410	2	H97639	probable D-amino a
747	29	51.8	293	2	T43586	secretion control	820	29	51.8	411	2	AB2214	glutamate-1-semial
748	29	51.8	293	2	S15320	yopN protein - Var	821	29	51.8	411	2	E71873	hypothetical prote
749	29	51.8	293	2	I40537	glycine betaine AB	822	29	51.8	412	2	G70735	hypothetical prote
750	29	51.8	294	2	T31946	hypothetical prote	823	29	51.8	412	2	AC2152	dolichyl-phosphate
751	29	51.8	298	2	C87535	hypothetical prote	824	29	51.8	413	2	T31663	isoleucine-tRNA li
752	29	51.8	298	2	G72531	hypothetical prote	825	29	51.8	414	2	AG2321	dolichol-phosphate
753	29	51.8	302	2	G70778	[acyl-carrier-prot	826	29	51.8	414	2	S06729	probable IS117 tra
754	29	51.8	303	2	E95374	probable ABC trans	827	29	51.8	418	2	T15142	hypothetical prote
755	29	51.8	303	2	C97463	hypothetical prote	828	29	51.8	419	2	S54102	isopenicillin N ep
756	29	51.8	304	2	F84169	hypothetical prote	829	29	51.8	420	2	A10921	UDP-MannNac dehydro
757	29	51.8	309	2	I53384	4-1BB ligand - mou	830	29	51.8	420	2	S12484	transforming prote
758	29	51.8	311	2	T52312	deacetoxycephalosp	831	29	51.8	428	2	JC4601	hypothetical 48.2k
759	29	51.8	311	2	A33043	deacetoxycephalosp	832	29	51.8	430	2	B84619	probable serine ca

833	29	51.8	430	2	B70518	probable cina - My	906	29	51.8	574	2	C83725	endo-beta-1,4-gluc
834	29	51.8	433	2	H82075	pho4 family protei	907	29	51.8	574	2	A84782	hypothetical prote
835	29	51.8	434	2	G70011	conserved hypotet	908	29	51.8	575	2	T48224	probable homeodoma
836	29	51.8	435	2	G75366	glucose-1-phosphat	909	29	51.8	584	2	D84679	probable homeodoma
837	29	51.8	437	2	D84619	probable serine ca	910	29	51.8	592	2	B83231	probable short-cha
838	29	51.8	438	2	T35355	hypothetical prote	911	29	51.8	594	1	WZBE82	protein-serine/thr
839	29	51.8	441	2	D84511	En/spm-like transp	912	29	51.8	598	2	T02795	probable membrane
840	29	51.8	441	2	H82642	hypothetical prote	913	29	51.8	600	2	H81331	exonuclease ABC c
841	29	51.8	441	2	H87523	argininosuccinate	914	29	51.8	600	2	H98149	ATP-binding transp
842	29	51.8	442	2	T02620	hypothetical prote	915	29	51.8	600	2	AD3138	hypothetical prote
843	29	51.8	442	2	D95889	conserved hypotet	916	29	51.8	611	2	S09142	ND5 intron 1 prote
844	29	51.8	445	2	F75389	probable carboxyl-	917	29	51.8	627	2	T02415	probable homeodoma
845	29	51.8	446	2	F82609	aminopeptidase P X	918	29	51.8	629	2	AE2497	hypothetical prote
846	29	51.8	446	2	AG3284	high-affinity bran	919	29	51.8	638	2	D85435	BEU1-like homeobox
847	29	51.8	447	2	G96569	hypothetical prote	920	29	51.8	638	2	H82690	hypothetical prote
848	29	51.8	447	2	AF3459	hypothetical prote	921	29	51.8	648	2	B40727	hypothetical prote
849	29	51.8	447	2	A13364	hypothetical prote	922	29	51.8	650	2	AG0733	probable bacteriop
850	29	51.8	449	2	F84241	hypothetical prote	923	29	51.8	655	1	A46698	probable growth
851	29	51.8	450	2	AG0008	probable membrane	924	29	51.8	655	2	E82712	hepatocyte growth
852	29	51.8	451	2	B96495	hypothetical prote	925	29	51.8	660	2	AH2348	probable carboxydr
853	29	51.8	453	2	E69755	hypothetical prote	926	29	51.8	660	2	A69138	DNA-dependent RNA
854	29	51.8	455	2	H84935	UDP-N-acetylmuram	927	29	51.8	664	2	A69138	flagellar hook-ass
855	29	51.8	456	2	A70920	probable adrenodox	928	29	51.8	666	2	G82111	galactosylceramida
856	29	51.8	463	2	G84901	hypothetical prote	929	29	51.8	666	2	F84139	guanosine-3',5'-bi
857	29	51.8	464	2	T17332	hypothetical prote	930	29	51.8	677	2	F84139	probable FMN oxido
858	29	51.8	466	2	AE2999	argininosuccinate	931	29	51.8	686	2	F82971	probable FMN oxido
859	29	51.8	467	2	D83055	poly(A) polymerase	932	29	51.8	690	2	H87691	penicillin-binding
860	29	51.8	469	1	NM1V27	exo-alpha-sialidas	933	29	51.8	690	2	A84639	probable glucosylt
861	29	51.8	469	1	NM1VH3	exo-alpha-sialidas	934	29	51.8	692	2	T00025	PSB-95 binding pro
862	29	51.8	469	1	NM1VN2	exo-alpha-sialidas	935	29	51.8	692	2	G95844	probable pyruvate
863	29	51.8	469	1	NM1V2	exo-alpha-sialidas	936	29	51.8	695	2	S62400	probable pyruvate
864	29	51.8	469	2	JQ1644	hypothetical prote	937	29	51.8	702	1	SHECGD	amphiphysin (clone
865	29	51.8	469	2	D82976	probable ferredoxi	938	29	51.8	702	2	E91194	guanosine-3',5'-bi
866	29	51.8	479	2	C75313	ribosomal protein	939	29	51.8	702	2	F86041	(p)ppopp synthetas
867	29	51.8	480	2	S56639	probable flavonol	940	29	51.8	702	2	AG0005	guanosine-3',5'-bi
868	29	51.8	481	2	B84700	ribosomal protein	941	29	51.8	703	2	AG0969	guanosine-3',5'-bi
869	29	51.8	483	2	C82090	argininosuccinate	942	29	51.8	709	2	T32089	hypothetical prote
870	29	51.8	486	2	D98284	probable serine-ty	943	29	51.8	718	2	E82697	GRP diphosphokinas
871	29	51.8	490	2	B71279	hypothetical prote	944	29	51.8	720	2	T51007	hypothetical prote
872	29	51.8	493	2	A12314	transcription term	945	29	51.8	731	2	JC7701	ARHGAP9 protein -
873	29	51.8	495	1	FJEC	transcription term	946	29	51.8	769	2	G83183	probable chemotaxi
874	29	51.8	495	2	B91135	transcription term	947	29	51.8	781	2	A43866	neuraminidase - Vi
875	29	51.8	495	2	H85980	transcription term	948	29	51.8	783	2	T45899	receptor protein k
876	29	51.8	496	2	H83577	probable transcrip	949	29	51.8	787	2	A75347	GRP pyrophosphokin
877	29	51.8	497	2	T48730	related to PR9 pr	950	29	51.8	796	2	T34805	hypothetical prote
878	29	51.8	500	1	A53377	transcription term	951	29	51.8	807	2	E82158	neuraminidase VC17
879	29	51.8	503	2	T37119	probable membrane-	952	29	51.8	814	2	G02390	disintegrin-like m
880	29	51.8	505	2	H95390	probable reverse t	953	29	51.8	821	2	S39983	ep88 protein - mou
881	29	51.8	506	2	B84339	hypothetical prote	954	29	51.8	822	2	I38728	epidermal growth f
882	29	51.8	507	2	A25228	protein kinase CDC	955	29	51.8	840	2	AG0526	penicillin-binding
883	29	51.8	509	2	S49224	glycoprotein precu	956	29	51.8	842	2	C83177	probable phosphotr
884	29	51.8	510	2	I56242	lymphoid cell acti	957	29	51.8	844	1	ZPECPB	penicillin-binding
885	29	51.8	511	2	S05081	hypothetical prote	958	29	51.8	844	2	A85499	peptidoglycan synt
886	29	51.8	512	2	H84310	cobyrlic acid synth	959	29	51.8	844	2	A99648	peptidoglycan synt
887	29	51.8	512	2	G85916	hypothetical prote	960	29	51.8	846	2	S57580	pepidoglycan synt
888	29	51.8	512	2	D91072	multidrug resistan	961	29	51.8	851	2	T47495	hypothetical prote
889	29	51.8	512	2	G65048	multidrug resistan	962	29	51.8	868	2	AF3204	autotransporter pr
890	29	51.8	512	2	AC3399	probable colanic b	963	29	51.8	880	1	B33926	DNA-directed RNA p
891	29	51.8	514	2	S52771	beta-glucosidase (964	29	51.8	880	2	G90163	sensor protein Kdp
892	29	51.8	514	2	F81811	conserved hypotet	965	29	51.8	902	2	AD0587	assimilatory nitra
893	29	51.8	524	2	F96784	hypothetical prote	966	29	51.8	908	2	A83424	penicillin-binding
894	29	51.8	525	2	H70982	probable fadD7 pro	967	29	51.8	911	2	H97053	iodide peroxidase
895	29	51.8	525	2	E87076	probable secreted	968	29	51.8	926	1	OPFGIT	hypothetical prote
896	29	51.8	528	2	C96564	probable beta-gluc	969	29	51.8	957	2	H69141	isoleucyl-tRNA syn
897	29	51.8	528	2	D70968	hypothetical prote	970	29	51.8	967	2	AH2660	isoleucyl-tRNA syn
898	29	51.8	530	2	JN0597	calnexin-like prot	971	29	51.8	967	2	F97442	hypothetical protei
899	29	51.8	532	2	T49873	calnexin homolog -	972	29	51.8	968	2	T45746	DAP-1-alpha protei
900	29	51.8	538	2	H86329	PeF9.25 protein -	973	29	51.8	977	2	T00014	Ca2+-transporting
901	29	51.8	548	2	D87254	N utilization subs	974	29	51.8	994	2	S24359	hypothetical prote
902	29	51.8	556	2	T47552	hypothetical prote	975	29	51.8	1025	2	AC0930	hypothetical prote
903	29	51.8	561	2	T46845	K+-transporting Ar	976	29	51.8	1025	2	AB0836	probable bacteriop
904	29	51.8	564	2	C84188	hypothetical prote	977	29	51.8	1034	2	AB0551	exonuclease SbcC l
905	29	51.8	573	2	B70942	hypothetical prote	978	29	51.8				

979 29 51.8 1045 2 DB3393
980 29 51.8 1046 2 S67786
981 29 51.8 1051 2 T18302
982 29 51.8 1054 2 T30901
983 29 51.8 1064 2 H97657
984 29 51.8 1064 2 AH2881
985 29 51.8 1113 2 T00271
986 29 51.8 1114 2 T49517
987 29 51.8 1128 2 AH1949
988 29 51.8 1154 2 T48829
989 29 51.8 1194 2 E96624
990 29 51.8 1200 2 C96025
991 29 51.8 1228 2 AE2397
992 29 51.8 1230 2 T04181
993 29 51.8 1231 2 C84716
994 29 51.8 1254 2 G86379
995 29 51.8 1263 2 T19472
996 29 51.8 1271 2 S37958
997 29 51.8 1474 2 B85188
998 29 51.8 1485 2 AG0171
999 29 51.8 1490 2 T30550
1000 29 51.8 1505 2 JC4851

ALIGNMENTS

RESULT 1
S76487
probable copper-transporting ATPase (EC 3.6.1.-) - Synecchocystis sp. (strain PCC 6803)
C:Species: Synecchocystis sp.
A:Variety: PCC 6803
C:Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 09-Jul-2004
A:Accession: S76487
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-780 <XAN>
A:Cross-references: UNIPROT:P74512; UNIPARC:UPI00000D35A0; EMBL:D90915; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Function:

A:Description: copper transport coupled with ATP hydrolysis
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
C:Keywords: ATP; enterococcus; copper transport; hydrolase; ion transport; phosphoprotein; transmembra
F:30-59/Domain: heavy-metal-associated homology <HMA>
F:149-505/Domain: ATPase transduction domain homology <ATT>
F:581-725/Domain: ATPase nucleotide-binding domain homology <ATN>
F:35-38/Binding site: copper (Cys) #status predicted
F:302/Active site: Glu #status predicted
F:460/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 73.2%; Score 41; DB 1; Length 780;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|||||
Db 357 RLADQVAGW 365

RESULT 2
E87684
ATP phosphoribosyltransferase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
A:Accession: E87684
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87684
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <STO>
A:Cross-references: UNIPROT:Q9A2P5; UNIPARC:UPI000012C73E; GB:AE005673; NID:G13425239; P
C:Genetics:
A:Gene: CC3511

Query Match 69.6%; Score 39; DB 2; Length 320;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|||||
Db 13 GRLKQVAGW 22

RESULT 3

AD1929
hypothetical protein all0983 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
A:Accession: AD1929
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD1929
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <KUR>
A:Cross-references: UNIPROT:Q8YV68; UNIPARC:UPI00000CDF37; GB:BA000019; PIDN:BA872940.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0983
C:Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 69.6%; Score 39; DB 2; Length 360;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|||||
Db 265 GHVKDQYTGW 274

RESULT 4

T34101

hypothetical protein, neural-specific - mouse

C:Species: Mus musculus (house mouse)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
A:Accession: T34101; T14159
R:Bosolasco, M.; Croci, L.; Corradi, A.; Consalez, G.G.
submitted to the EMBL Data Library, June 1996

A:Reference number: Z21477
A:Accession: T34101
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1216 <BOS>

A:Cross-references: UNIPROT:Q9QW16; UNIPARC:UPI0000029B36; EMBL:U59873; NID:G4097422; PI
R:Croci, L.; Bosolasco, M.; Consalez, G.G.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z17896
A:Accession: T14159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A;Residues: 1-1086,'KL',1089-1216 <CRO>
A;Cross-references: UNIPARC:UPI0000027C2C; EMBL:AF040944; NID:g3098417; PID:g3098418; PI
C;Genetics:
A;Map position: 11
A;Note: P140

Query Match 69.6%; Score 39; DB 2; Length 1216;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
||:||||
Db 104 RMREQVGGW 112

RESULT 5
B70725
probable linB protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70725
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70725
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-300 <COL>
A;Cross-references: UNIPROT:Q50642; UNIPARC:UPI000012926D; GB:Z77724; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: linB
C;Superfamily: tropinesterase

Query Match 67.9%; Score 38; DB 2; Length 300;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 252 GRIRDYVRWSW 261

RESULT 6
AB2809
conserved hypothetical protein Atul892 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AB2809
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB2809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <KUR>
A;Cross-references: UNIPROT:Q8UE69; UNIPARC:UPI00000D1CF3; GB:AE008688; PIDN:AAL42888.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul892
A;Map position: circular chromosome

Query Match 66.1%; Score 37; DB 2; Length 63;
Best Local Similarity 50.0%; Pred. No. 5;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
||:||||
Db 22 GRTLDEIGGW 31

RESULT 7

C94233

hypothetical protein tfeA [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: C94233

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabc

Jung, K.H.; Alam, M.; Freitas, T.

Proc Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A;Title: Genome sequence of Halobacterium species NRC-1

A;Reference number: A94160; MUID:20504483; PMID:11016950

A;Accession: C94233

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-181 <STO>

A;Cross-references: UNIPROT:Q9HRC7; UNIPARC:UPI0000063708; GB:AE004437; NID:g10580335; I

C;Genetics:

A;Gene: tfeA

C;Superfamily: hypothetical protein MJ0777

Query Match 66.1%; Score 37; DB 2; Length 181;

Best Local Similarity 66.7%; Pred. No. 15;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10

||||:|

Db 78 RVRDEDSGW 86

RESULT 8

S02828

albicidin resistance protein - Klebsiella oxytoca

C;Species: Klebsiella oxytoca

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: S02828

R;Walker, M.J.; Birch, R.G.; Pemberton, J.M.

Mol. Microbiol. 2, 443-454, 1988

A;Title: Cloning and characterization of an albicidin resistance gene from Klebsiella o

A;Reference number: S02828; MUID:89013885; PMID:2845223

A;Accession: S02828

A;Molecule type: DNA

A;Residues: 1-218 <WAL>

A;Cross-references: UNIPROT:PI0488; UNIPARC:UPI00001257BF; EMBL:Y00558; NID:g43785; PID

A;Note: Part of this sequence, including the amino end of the mature protein, was confi.

C;Comment: This protein forms an inactive complex with albicidin.

C;Superfamily: Klebsiella oxytoca albicidin resistance protein

C;Keywords: antibiotic resistance

F;1-218/Product: albicidin resistance protein #status experimental <MAT>

Query Match 66.1%; Score 37; DB 1; Length 218;

Best Local Similarity 70.0%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10

||||:|

Db 48 GLRLWQPAGW 57

RESULT 9

C95880

probable protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasamid pSymb

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C;Accession: C95880

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 PROC. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A;Reference number: A95842; MUID:21396508; PMID:11481431
 A;Accession: C95880
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-421 <KUR>
 A;Cross-references: UNIPROT:Q92WM5; UNIPARC:UPI00000CBAAA; GB:AL591985; PIDN:CAC48707.1;
 A;Experimental source: strain 1021, megaplasmid pSymb
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pala, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics:
 A;Gene: SMB20321
 A;Genome: plasmid
 C;Superfamily: conserved hypothetical protein HI1029

Query Match 66.1%; Score 37; DB 2; Length 421;
 Best Local Similarity 70.0%; Pred. No. 36;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
 | | | | |
 DB 293 GNVPPQVAGW 302

RESULT 10
 A12146
 excinuclease ABC chain C uvrC [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: A12146
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: A12146
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-627 <KUR>
 A;Cross-references: UNIPROT:Q8YTI9; UNIPARC:UPI0000137ED4; GB:BA0000019; PIDN:BAB74427.1;
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: uvrC
 C;Superfamily: excinuclease ABC chain C

Query Match 66.1%; Score 37; DB 2; Length 627;
 Best Local Similarity 75.0%; Pred. No. 54;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVRDQVAG 9
 | | | | |
 DB 239 RIRDQIAG 246

RESULT 11
 S36741
 probable copper-transporting ATPase (EC 3.6.1.-) pacs - Synecococcus sp.
 C;Species: Synecococcus sp.
 C;Date: 16-Oct-1998 #sequence_revision 16-Oct-1998 #text_change 16-Jun-2000
 A;Accession: S36741
 R;Kanamaru, K.; Kashiwagi, S.; Mizuno, T.
 FEBS Lett. 330, 99-104, 1993
 A;Title: The cyanobacterium, Synecococcus sp. PCC7942, possesses two distinct genes end

A;Reference number: S36741; MUID:93380581; PMID:8370468
 A;Accession: S36741
 A;Molecule type: DNA
 A;Residues: 1-747 <KAN>
 A;Cross-references: UNIPARC:UPI000012620F; GB:D16437; MID:9435124; PIDN:BAA03907.1; PID:
 C;Function:
 A;Description: copper transport coupled with ATP hydrolysis
 C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d
 C;Keywords: ATP; copper transport; hydrolase; ion transport; phosphoprotein; transmembra
 F;9-38/Domain: heavy-metal-associated homology <HMA>
 F;136-480/Domain: ATPase transduction domain homology <ATT>
 F;557-699/Domain: ATPase nucleotide-binding domain homology <ATN>
 F;14.17/Binding site: copper (Cys) #status predicted
 F;289/Active site: Glu #status predicted
 F;43/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 66.1%; Score 37; DB 1; Length 747;
 Best Local Similarity 66.7%; Pred. No. 65;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
 | | | | |
 DB 344 RLADQVTGW 352

RESULT 12
 AH2543
 cation-transporting ATPase alr7635 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AH2543
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AH2543
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-753 <KUR>
 A;Cross-references: UNIPROT:Q8ZS77; UNIPARC:UPI00000CCDC5; GB:AP003602; PIDN:BA077278.1;
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr7635
 A;Genome: plasmid
 C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d

Query Match 66.1%; Score 37; DB 2; Length 753;
 Best Local Similarity 66.7%; Pred. No. 66;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
 | | | | |
 DB 345 RLADQVTGW 353

RESULT 13
 AE2009
 cation-transporting ATPase alr1627 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AE2009
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AE2009
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-753 <KUR>

A;Cross-references: UNIPROT:Q8YWI6; UNIPARC:UPI00000CE159; GB:BA000019; PIDN:BA077993.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1627
C;Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding d

Query Match 66.1%; Score 37; DB 2; Length 753;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: ||| ||
Db 345 RLADQVTGW 353

RESULT 14
A45407
collagen alpha 3(IV) chain - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 22-Sep-1993 #sequence revision 18-Nov-1994 #text_change 31-Dec-2004
C;Accession: A45407; A43903; A23940
R;Exposito, J.Y.; D'Alessio, M.; Di Liberto, M.; Ramirez, F.
J. Biol. Chem. 268, 5249-5254, 1993
A;Title: Complete primary structure of a sea urchin type IV collagen alpha chain and ana
A;Reference number: A45407; MUID:93186842; PMID:8444899
A;Accession: A45407
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1752 <EXP>
A;Cross-references: UNIPROT:Q26312; UNIPARC:UPI00001773DE
A;Note: sequence extracted from NCBI backbone (NCBIP:126841)
R;Wessel, G.M.; Etkin, M.; Benson, S.
Dev. Biol. 148, 261-272, 1991
A;Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ
A;Reference number: A43903; MUID:92038439; PMID:1936564
A;Accession: A43903
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'P', 633-1537, 'G' <WES>
A;Cross-references: UNIPARC:UPI000007C802; GB:S64572; NID:9238616; PIDN:AAB20270.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBIP:64573)
R;Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986
A;Title: Structure and developmentally regulated expression of a Strongylocentrotus purp
A;Reference number: A23940; MUID:86205894; PMID:3458186
A;Accession: A23940
A;Molecule type: DNA
A;Residues: 742-812 <VEN>
A;Cross-references: UNIPARC:UPI00001773DF; EMBL:M13206
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;29-161/Domain: amino-terminal nonhelical, 7S <7SD>
F;162-1523/Region: interrupted helical
F;1524-1752/Domain: carboxyl-terminal nonhelical, NCI <NCI>
F;1534-1634/Domain: collagen IV carboxyl-terminal repeat <CTI>
F;1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT>
F;129/Modified site: allylsine (lys) #status predicted

Query Match 66.1%; Score 37; DB 2; Length 1752;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|: ||| ||
Db 1165 GIIRDLLAGW 1174

RESULT 15
G82724
hypothetical protein XF1092 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82724
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below

A;Accession: G82724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <SIM>
A;Cross-references: UNIPROT:Q9PED6; UNIPARC:UPI00000C2605; GB:AE003845; GB:AE003849; NII
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; f
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carrer, i
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, f
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; i
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1092

Query Match 64.3%; Score 36; DB 2; Length 67;
Best Local Similarity 55.6%; Pred. No. 8.3;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: ||| ||
Db 20 RIRDETASW 28

RESULT 16
G82860
replication protein XPB0001 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82860
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below

A;Accession: G82860
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-283 <SIM>
A;Cross-references: UNIPROT:Q9PHK6; UNIPARC:UPI00000C223B; GB:AE003850; NID:99112233; P
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; f
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, i
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, f
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; i
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XPB0001
A;Genome: plasmid
A;Note: plasmid pXF1.3
C;Superfamily: class II filamentous phage gene 278 protein

```

Query Match      64.3%; Score 36; DB 2; Length 283;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
   ||:|||||
Db 143 GRIDGVVGW 152

RESULT 17
C75353
ribosomal large subunit pseudouridine synthase D - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: C75353
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036996; PMID:10567266
A:Accession: C75353
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <WHI>
A:Cross-references: UNIPROT:Q9RTH5; UNIPARC:UPI0000003BD4; GB:AE002021; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1789
A:Map position: 1
C:Superfamily: conserved hypothetical protein HI0176

Query Match      64.3%; Score 36; DB 2; Length 338;
Best Local Similarity 70.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
   ||:|||||
Db 42 GVSRSQVAGW 51

RESULT 18
E83416
hypothetical protein PA1829 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: E83416
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83416
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <STO>
A:Cross-references: UNIPROT:Q912R6; UNIPARC:UPI000000C54B9; GB:AE004609; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1829

Query Match      64.3%; Score 36; DB 2; Length 356;
Best Local Similarity 60.0%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
   ||:|||||
Db 168 GYVQRQIAGW 177

RESULT 19
F70686

```

```

hypothetical protein Rv1723 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70686
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70686
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-415 <COL>
A:Cross-references: UNIPROT:P71981; UNIPARC:UPI000000D5F1E; GB:Z81360; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1723

Query Match      64.3%; Score 36; DB 2; Length 415;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
   ||:|||||
Db 174 VREQVIGW 181

RESULT 20
G86207
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G86207
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86207
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <STO>
A:Cross-references: UNIPROT:Q91ML6; UNIPARC:UPI000000BD55; GB:AE005172; NID:gl0567858; P
C:Genetics:
A:Map position: 1
C:Superfamily: flavonol O3-glucosyltransferase

Query Match      64.3%; Score 36; DB 2; Length 479;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
   ||:|||||
Db 269 RDQIVGW 275

RESULT 21
I39066
N-methyl-D-aspartate receptor chain NMDAR2B - human (fragments)
N:Alternate names: glutamate receptor, ionotropic, N-methyl D-aspartate 2B; NMDA receptor
C:Species: Homo sapiens (man)
C:Date: 01-Mar-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C:Accession: I39066; I39067; I39064
R:Mandich, P.; Schito, A.M.; Bellone, E.; Antonacci, R.; Finelli, P.; Rocchi, M.; Ajmar,
Genomics 22, 216-218, 1994

```

A>Title: Mapping of the human NMDAR2B receptor subunit gene (GRIN2B) to chromosome 12p12
A:Reference number: I39064; MUID:95048375; PMID:7959773
A:Accession: I39066
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-294 <RES>
A:Cross-references: UNIPROT:Q13224; UNIPARC:UPI000016A1FC; EMBL:U28861; NID:G899434; PID:PI
A:Accession: I39067
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 295-723 <RE2>
A:Cross-references: UNIPARC:UPI000016A1FD; EMBL:U28862; NID:G899436; PIDN:AAA69920.1; PI
A:Accession: I39064
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 395-559, 'R', 561-613 <RE3>
A:Cross-references: UNIPARC:UPI000016A1F9; EMBL:U28758; NID:G965063; PIDN:AAA74930.1; PI
C:Genetics:
A:Gene: GDB:GRIN2B; NMDAR2B
A:Cross-references: GDB:134710; OMIM:138252
A:Map position: 12p12-12p12
C:Superfamily: N-methyl-D-aspartate receptor 2A; glutamate receptor homology
C:Keywords: neurotransmitter receptor

Query Match 64.3%; Score 36; DB 2; Length 723;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVRDQVAG 9
|||||:
DB 292 RVRDQVSG 299

RESULT 22
G82965
conserved hypothetical protein PA5442 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Aug-2004
C:Accession: G82965
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G82965
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-951 <STO>
A:Cross-references: UNIPROT:Q9HTC4; UNIPARC:UPI000000C600B; GB:AE004957; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5442

Query Match 64.3%; Score 36; DB 2; Length 951;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAG 10
|||||:
DB 354 RVRDREGW 362

RESULT 23
GNLR11
retrovirus-related reverse transcriptase pseudogene - slow loris
C:Species: Nycticebus coucang (slow loris)
C>Date: 31-Mar-1998 #sequence_revision 04-Jan-1996 #text_change 09-Jul-2004
C:Accession: B25313
R:Hattori, M.; Kuhara, S.; Takenaka, O.; Sakaki, Y.
Nature 321, 625-628, 1986
A>Title: L1 family of repetitive DNA sequences in primates may be derived from a sequenc
A:Reference number: A93381; MUID:86230917; PMID:2423883

A:Accession: B25313
A>Status: conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-1260 <HAT>
A:Cross-references: UNIPROT:P08548; UNIPARC:UPI0000012E69D
A>Note: this sequence was constructed from an alignment of six sequences, determined by
C:Keywords: reverse transcriptase; pseudogene

Query Match 64.3%; Score 36; DB 4; Length 1260;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAG 10
|||||:
DB 1063 RTRDPISGW 1071

RESULT 24
S58380
probable RNA-directed DNA polymerase (EC 2.7.7.49) - Crithidia fasciculata retrotranspo
N:Alternate names: probable reverse transcriptase
C:Species: Crithidia fasciculata
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Sep-1997
A:Accession: S58380; S58381
R:Gabriel, A.
submitted to the EMBL Data Library, December 1994
A:Reference number: S58380
A:Accession: S58380
A:Molecule type: DNA
A:Residues: 1-2517 <GAB>
A:Cross-references: UNIPARC:UPI000011DBA6; EMBL:U19151; NID:G624680; PID:G624681
R:Teng, S.C.; Wang, S.X.; Gabriel, A.
Nucleic Acids Res. 23, 2929-2936, 1995
A>Title: A new non-LTR retrotransposon provides evidence for multiple distinct site-spe
A:Reference number: S58381; MUID:95388509; PMID:7659515
A:Accession: S58381
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1825-2116 <TEN>
A:Cross-references: UNIPARC:UPI000017B579; EMBL:U19151
A:Experimental source: retrotransposon CRE2
C:Genetics:
A:Mobile element: retrotransposon CRE2
C:Keywords: nucleotidyltransferase

Query Match 64.3%; Score 36; DB 2; Length 2517;
Best Local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAG 10
|||||:
DB 2351 KVRDVTWIGW 2359

RESULT 25
G82245
transcription regulator ArpR family VC1068 [imported] - Vibrio cholerae (strain N16961
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82245
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82245
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <HEI>
A:Cross-references: UNIPROT:Q9KT37; UNIPARC:UPI000000C2E7D; GB:AE004188; GB:AE003852; NT
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC1068
A:Map position: 1

Query Match 62.5%; Score 35; DB 2; Length 113;
Best Local Similarity 55.6%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:
Db 69 RISDQLPGW 77

RESULT 26

B69047
arsenate reductase - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69047
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Ji, Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; S.; Church, G.M.; Daniele, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func0
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69047
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <MTH>
A:Cross-references: UNIPROT:Q27408; UNIPARC:UPI000006656D; GB:AE000898; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1355
C:Superfamily: protein-tyrosine-phosphatase, low molecular weight

Query Match 62.5%; Score 35; DB 2; Length 130;
Best Local Similarity 44.4%; Pred. No. 25;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:
Db 112 RIRDSISDW 120

RESULT 27

A41840
hypothetical protein (hadA 5' region) - Pseudomonas putida (fragment)
C:Species: Pseudomonas putida
C:Date: 04-Mar-1993 #sequence_revision 26-Apr-1996 #text_change 21-Aug-1998
C:Accession: A41840
R:Barth, P.T.; Bolton, L.; Thomson, J.C.
J. Bacteriol. 174, 2612-2619, 1992
A:Title: Cloning and partial sequencing of an operon encoding two Pseudomonas putida hal
A:Reference number: A41840; MUID:92210506; PMID:1556080
A:Contents: AJ1
A:Accession: A41840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <BAR>
A:Cross-references: UNIPARC:UPI00000178057
A:Note: sequence extracted from NCBI backbone (NCBIN:94091, NCBIP:94092)
C:Superfamily: proline carrier protein

Query Match 62.5%; Score 35; DB 2; Length 150;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:
Db 118 RVRGLAGW 126

RESULT 28

AD2083

phosphonate ABC transporter, ATP-binding component phnK [imported] - Nostoc sp. (strain
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004
C:Accession: AD2083
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2083
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <KUR>
A:Cross-references: UNIPROT:Q8YUW3; UNIPARC:UPI000000CE374; GB:BA000019; PIDN:BAB73917.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: phnK

Query Match 62.5%; Score 35; DB 2; Length 257;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|:|:
Db 124 GNRDEAAHW 133

RESULT 29

S76855
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76855
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76855
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <KAN>
A:Cross-references: UNIPROT:P74651; UNIPARC:UPI00000C09EB; EMBL:D90917; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 62.5%; Score 35; DB 2; Length 300;
Best Local Similarity 55.6%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|:|:
Db 206 RYKDEVVGM 214

RESULT 30

B87128
conserved hypothetical protein ML1752 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: B87128
R:Cole, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squires, R.; Sq

A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: B87128
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-302 <STO>
A:Cross-references: UNIPROT:Q9CBP4; UNIPARC:UPI00000C6DB1; GB:AL450380; NID:GL3093491; H
C:Genetics:
A:Gene: ML1752

Query Match 62.5%; Score 35; DB 2; Length 302;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVDRQVAGW 10
| : | : | : |
Db 27 GEVKDRSSGW 36

RESULT 31
G83219
probable dihydrolipoamide acetyltransferase PA3415 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: G83219
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83219
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <STO>
A:Cross-references: UNIPROT:Q9HYJ0; UNIPARC:UPI00000C59FD; GB:AE004762; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3415

Query Match 62.5%; Score 35; DB 2; Length 370;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVDRQVAGW 10
| : | : | : |
Db 319 GGRDRVWAW 328

RESULT 32
S38461
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Chlorobium limicola
C:Species: Chlorobium limicola
C:Date: 22-Jan-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S38461
R:Schuetz, M.; Zirnigbl, S.; Le Coutre, J.; Buettner, M.; Xie, D.L.; Nelson, N.; Deutzma
submitted to the EMBL Data Library, June 1993
A:Description: A transcription unit for the Rieske FeS-protein and cytochrome b in Chlor
A:Reference number: S38460
A:Accession: S38461
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <SCH>
A:Cross-references: UNIPROT:Q59297; UNIPARC:UPI000016BB31; EMBL:X73628; NID:G410433; PII
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase
F;89-423/Domain: cytochrome b homology <CBH>
F;89-289/Domain: cytochrome b6 homology <CB6>
F;304-423/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;161,261/Binding site: heme iron (His) (axial ligands) (low potential) #status predict
F;175,276/Binding site: heme iron (His) (axial ligands) (high potential) #status predict

Query Match 62.5%; Score 35; DB 2; Length 428;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDOVAGW 10
| : | : | : |

Db 74 RDAVAGW 80

RESULT 33
G84823
probable anthocyanin 5-aromatic acyltransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: G84823
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84823
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <STO>
A:Cross-references: UNIPROT:O04201; UNIPARC:UPI00000396F7; GB:AE002093; NID:G2088651; P1
C:Genetics:
A:Gene: At2G39980
A:Map position: 2
C:Superfamily: N-hydroxycinnamoyl/benzoyl transferase

Query Match 62.5%; Score 35; DB 2; Length 482;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVDRQVAGW 10
| : | : | : |
Db 382 GRIRSVADW 391

RESULT 34
E95335
probable cation transport P-type ATPase (EC 3.6.3.-) [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: E95335
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9886, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: E95335
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-733 <KUR>
A:Cross-references: UNIPROT:Q92ZA2; UNIPARC:UPI00000CB13D; GB:AE006469; PIDN:AAK65247.1
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
heault, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma1087
A:Gene: plasmid
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding
C:Keywords: hydrolase

Query Match 62.5%; Score 35; DB 2; Length 733;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDRQVAGW 10
| : | : | : |
Db 330 RLADVSGW 338

```
RESULT 35
H75507
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: H75507
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75507
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-891 <WHI>
A:Cross-references: UNIPROT:Q9RMV5; UNIPARC:UPI00000C1789; GB:AE001911; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0530
A:Map position: 1

Query Match 62.5%; Score 35; DB 2; Length 891;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
|||: ||
Db 566 VRDELTGW 573

RESULT 36
C72409
reverse gyrase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
A:Accession: C72409
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72409
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1104 <ARN>
A:Cross-references: UNIPROT:O51934; UNIPARC:UPI00000D382F; GB:AE001702; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0173

Query Match 62.5%; Score 35; DB 2; Length 1104;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||: ||
Db 752 GRVQSTVLGW 761

RESULT 37
E72611
probable ATP-dependent DNA helicase APE1353 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: E72611
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
```

```
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72611
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1273 <RAW>
A:Cross-references: UNIPROT:Q9YCAL; UNIPARC:UPI000005DF8; DBJ:AF000061; NID:G5104821;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1353

Query Match 62.5%; Score 35; DB 2; Length 1273;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||: ||
Db 341 GLRIRVRGW 350

RESULT 38
AC3434
calf thymus ribonuclease H (EC 3.1.26.4) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AC3434
R:DeIvecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujfer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goleman, E.; Sekov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3434
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <KUR>
A:Cross-references: UNIPROT:Q8YFR3; UNIPARC:UPI0000058097; GB:AE008917; PIDN:AAL52638.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11457
A:Map position: 1
C:Superfamily: ribonuclease H

Query Match 60.7%; Score 34; DB 2; Length 154;
Best Local Similarity 62.5%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VRDQVAGW 10
|||: ||
Db 73 VRDGISGW 80

RESULT 39
C72782
hypothetical protein APE0244 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: C72782
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <KAW>
A:Cross-references: UNIPROT:Q9YFK4; UNIPARC:UPI000005DA75; DBJ:AF000058; NID:G5103389;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0244
C:Superfamily: conserved hypothetical protein H11648

Query Match 60.7%; Score 34; DB 2; Length 186;
Best Local Similarity 66.7%; Pred. No. 57;
```

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVRDQVAG 9
|:|:|:|
Db 75 GKVRDKVWG 83

RESULT 40
D83620
hypothetical protein PA0201 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83620
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: D83620
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <STO>
A:Cross-references: UNIPROT:Q9IGT7; UNIPARC:UPI00000C4F91; GB:AE004458; GB:AE004091; NID:10580297; F87259
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0201

Query Match 60.7%; Score 34; DB 2; Length 193;
Best Local Similarity 87.5%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAG 9
|:|:|:|
Db 86 RVRDQVRG 93

RESULT 41
G87457
guanylate kinase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87457
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: G87457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:Q9A7N9; UNIPARC:UPI000012DDDB; GB:AE005673; NID:g13423091; F87259
C:Genetics:
A:Gene: CCI681
C:Superfamily: guanylate kinase; guanylate kinase homology

Query Match 60.7%; Score 34; DB 2; Length 213;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAG 10
|:|:|:|
Db 157 RANDEVAAG 165

RESULT 42
C84229
hypothetical protein Vng0718c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84229

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, J.; Laub, M.T.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: AB4160; MUID:20504483; PMID:11016950
A:Accession: C84229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <STO>
A:Cross-references: UNIPROT:Q9HRF9; UNIPARC:UPI00000636ED; GB:AE004437; NID:g10580297; F87259
C:Genetics:
A:Gene: VNG0718C

Query Match 60.7%; Score 34; DB 2; Length 225;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAG 10
|:|:|:|
Db 25 GRVLDVLRGW 34

RESULT 43
F87259
hypothetical protein CC0087 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F87259
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: F87259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <STO>
A:Cross-references: UNIPROT:Q9ABY3; UNIPARC:UPI00000C6F16; GB:AE005673; NID:g13421188; F87259
C:Genetics:
A:Gene: CC0087

Query Match 60.7%; Score 34; DB 2; Length 252;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RDQVAG 10
|:|:|:|
Db 185 KDQVVGW 191

RESULT 44
A13628
peptidylprolyl isomerase [EC 5.2.1.8] [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: A13628
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova, J.; Mazur, M.; Goltzman, E.; Salkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes, J.
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis.
A:Reference number: AD3252; PMID:11756688
A:Accession: A13628
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-271 <KUR>
A:Cross-references: UNIPROT:Q8YBES; UNIPARC:UPI00000586A7; GB:AE008918; PIDN:AAL54196.1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110954
A:Map position: 11

C;Keywords: cis-trans-isomerase

Query Match 60.7%; Score 34; DB 2; Length 271;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RVRDQVAGW 10
| : : : : :
DB 228 VADKIAGW 235

RESULT 45

T16323

hypothetical protein F41C6.4 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T16323

R;Geisael, C.

submitted to the EMBL Data Library, October 1995

A;Description: The sequence of *C. elegans* cosmid F41C6.

A;Reference number: Z18495

A;Accession: T16323

A;Status: preliminary; translated from GB/EMBL/DBSJ

A;Molecule type: DNA

A;Residues: 1-337 <GEI>

A;Cross-references: UNIPROT:Q20269; UNIPARC:UPI000007AD01; EMBL:U39745; NID:g1049470; P1

C;Genetics:

A;Gene: CESP:F41C6.4

A;Introns: 27/2; 105/2; 219/3

Query Match 60.7%; Score 34; DB 2; Length 337;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
: : : : :
DB 173 KLENIAGW 181

RESULT 46

G71499

probable D-Ala-D-Ala carboxypeptidase - *Chlamydia trachomatis* (serotype D, strain UW3/Cx)

C;Species: *Chlamydia trachomatis*

C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004

C;Accession: G71499

R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998

A;Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trach*

A;Reference number: A71570; MUID:99000809; PMID:9784136

A;Accession: G71499

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-343 <ARN>

A;Cross-references: UNIPROT:O84555; UNIPARC:UPI00000D335A; GB:AE001326; GB:AE001273; NID

A;Experimental source: serotype D, strain UW-3/Cx

C;Genetics:

A;Gene: dacC

Query Match 60.7%; Score 34; DB 2; Length 343;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
: : : : :
DB 120 KVKEVSGW 128

RESULT 47

G95297

probable hydrolasee [imported] - *Sinorhizobium meliloti* (strain 1021) megaplasmid pSymA

C;Species: *Sinorhizobium meliloti*

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 31-Dec-2004

C;Accession: G95297

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95297
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <KUR>
A;Cross-references: UNIPROT:Q930B1; UNIPARC:UPI00000CB048; GB:AE006469; PIDN:AAK64945.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0551
A;Genome: plasmid
C;Superfamily: metal-dependent hydrolase (amidohydrolase)
Query Match 60.7%; Score 34; DB 2; Length 434;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GRVRDQVAG 9
| : : : : :
DB 354 GQVRDPIAG 362
RESULT 48
C65203
argininosuccinate lyase (EC 4.3.2.1) - *Escherichia coli* (strain K-12)
N;Alternate names: argininosuccinase
C;Species: *Escherichia coli*
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C65203; J03033; I41139
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of *Escherichia coli* K-12.
A;Reference number: A64720; MUID:97428617; PMID:9278503
A;Accession: C65203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-457 <BLAT>
A;Cross-references: UNIPROT:P11447; UNIPARC:UPI0000125BF5; GB:AE000470; GB:U00096; NID:9
A;Experimental source: strain K-12, substrain MGL655
R;Parsot, C.; Boyen, A.; Cohen, G.N.; Glansdorff, N.
Gene 68, 275-283, 1988
A;Title: Nucleotide sequence of *Escherichia coli* argB and argC genes: comparison of N-ac
es.
A;Reference number: J03031; MUID:89121510; PMID:2851495
A;Accession: J03033
A;Molecule type: DNA
A;Residues: 1-51 <PAR>
A;Cross-references: UNIPARC:UPI000016BDF0; GB:M21446; NID:g145332; PIDN:AAA23479.1; PID:
R;Charlier, D.; Plette, J.; Glansdorff, N.
Nucleic Acids Res. 10, 5935-5948, 1982
A;Title: is3 can function as a mobile promoter in *e.coli*.
A;Reference number: I41137; MUID:83064529; PMID:6292860
A;Accession: I41139
A;Status: preliminary; translated from GB/EMBL/DBSJ
A;Molecule type: DNA
A;Residues: 1-27, 'G', 29-39, 'R', 41-44 <RES>
A;Cross-references: UNIPARC:UPI00000B08F9; GB:J01590; NID:g145350; PIDN:AAB59147.1; PID:
C;Comment: In arginine biosynthesis Glutamate is first converted to N-acetylglutamate wh
C;Comment: The latter two reactions are catalyzed by N-acetylglutamatekinase and N-acety
C;Genetics:

Search completed: May 12, 2006, 10:51:04
Job time : 43.2368 secs

A:Gene: argH
C:Superfamily: argininosuccinate lyase
C:Keywords: amidine-lyase; carbon-nitrogen lyase

Query Match 60.7%; Score 34; DB 2; Length 457;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
|:|:|
Db 209 REQLAGW 215

RESULT 49
F86087
argininosuccinate lyase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F86087
C:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480, MUID:21074935; PMID:11206551
A:Accession: F86087
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <STO>
A:Cross-references: UNIPROT:Q8X730; UNIPARC:UPI0000125EF3; GB:AE005174; NID:gl2518882; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: argH
C:Superfamily: argininosuccinate lyase

Query Match 60.7%; Score 34; DB 2; Length 457;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
|:|:|
Db 209 REQLAGW 215

RESULT 50
A98240
argininosuccinate lyase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A98240
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A98240
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <HAY>
A:Cross-references: UNIPROT:Q8X730; UNIPARC:UPI0000125EF3; GB:BA000007; PIDN:BAE38312.1;
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: ECs4889
C:Superfamily: argininosuccinate lyase

Query Match 60.7%; Score 34; DB 2; Length 457;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 RDQVAGW 10
|:|:|
Db 209 REQLAGW 215

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-35
Perfect score: 56
Sequence: 1 GRVRDQVAGW 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	76.8	242	Q4JY45	CORYNEBACTE
2	41	73.2	467	Q8RSP4	RHIZOBIUM 1
3	41	73.2	776	Q4TW05	ERYTHROBACT
4	41	73.2	780	P74512	SYNTHY3
5	41	73.2	823	Q89W17	BRADYRHOZOB
6	41	73.2	832	Q9KZY7	STREPTOMYCE
7	40	71.4	130	Q7VLA1	HAEMOPHILUS
8	40	71.4	234	Q9RL18	STREPTOMYCE
9	40	71.4	302	Q7RL87	PLASMODIUM
10	40	71.4	313	Q4Z1J9	PLASMODIUM
11	40	71.4	315	Q4XWR8	PLASMODIUM
12	40	71.4	320	Q812X0	PLAF7
13	40	71.4	369	Q7RXX6	NEUROSPORA
14	39	69.6	38	Q7TIZ4	9ZZZZ
15	39	69.6	124	Q9FRG1	ORYZA SATIV
16	39	69.6	320	HIS1	CAUCU
17	39	69.6	334	Q4LRX7	9BURK
18	39	69.6	360	Q8YV68	ANABAEANA SP
19	39	69.6	1250	SNIP	MOUSE
20	39	69.6	1370	Q4QD2	LEISHMANIA
21	39	69.6	1808	Q5LMT4	SILPO
22	38	67.9	167	Q8H1Z7	9ROSI
23	38	67.9	233	Q62FH7	BURMA
24	38	67.9	234	Q63VC8	BURPS
25	38	67.9	243	Q4WRP3	ASPFLU
26	38	67.9	300	DHAA	MYCTU
27	38	67.9	300	1	MYCTU
28	38	67.9	349	Q7UV53	RHOBA
29	38	67.9	367	Q4H705	9DEIO
30	38	67.9	367	Q89SP5	BRAJA
31	38	67.9	369	Q747F0	GEOSIL

32	38	67.9	385	2	Q9L144	STRCO
33	38	67.9	443	2	Q74SA4	YERPE
34	38	67.9	448	2	Q66HY4	BRARE
35	38	67.9	461	2	Q668C3	YERPS
36	38	67.9	588	2	Q7X541	9ACTO
37	38	67.9	941	2	Q869Y0	DICDI
38	38	67.9	941	2	Q553Y8	DICDI
39	37	66.1	63	2	Q8UE69	AGRT5
40	37	66.1	121	2	Q4LUS0	9BURK
41	37	66.1	178	2	Q606W1	METCA
42	37	66.1	181	2	Q9HRC7	HALSA
43	37	66.1	218	1	ALBR	KLEOX
44	37	66.1	244	2	Q67J54	SYMTH
45	37	66.1	276	2	Q5UQ11	MIMIV
46	37	66.1	307	2	Q4UPY2	XANCP
47	37	66.1	307	2	Q4NPES	9DELT
48	37	66.1	307	2	Q8P4D1	XANCP
49	37	66.1	335	2	Q4H433	SOLTU
50	37	66.1	372	2	Q5ON02	ENTHI
51	37	66.1	395	2	Q5QUR4	IDILO
52	37	66.1	420	2	Q8CEB1	MOUSE
53	37	66.1	421	2	Q92WM5	RHIME
54	37	66.1	431	2	Q7WLO2	BORBR
55	37	66.1	461	2	Q6ZE25	ORYSA
56	37	66.1	477	2	Q8RTW0	9PROT
57	37	66.1	477	2	Q6SFG3	9BACT
58	37	66.1	508	2	Q6DFS1	XENTR
59	37	66.1	508	2	Q6GP74	XENLA
60	37	66.1	510	1	ENP1	MOUSE
61	37	66.1	510	2	Q921Q6	MOUSE
62	37	66.1	510	2	Q544U5	MOUSE
63	37	66.1	510	2	Q6AZW0	BRARE
64	37	66.1	511	1	ENP1	RAT
65	37	66.1	539	2	Q8CDV7	MOUSE
66	37	66.1	579	2	Q65J87	BACLD
67	37	66.1	605	2	Q8NQY9	CORGL
68	37	66.1	621	2	Q82KB5	STRAM
69	37	66.1	627	1	UVRC	ANASP
70	37	66.1	627	2	Q5LPF8	SILPO
71	37	66.1	667	2	Q51TU2	MAGGR
72	37	66.1	679	2	Q5SMO4	ORYSA
73	37	66.1	747	1	ATCS	SYNP7
74	37	66.1	747	2	Q8RPV5	SYNP7
75	37	66.1	747	2	Q5N650	SYNP6
76	37	66.1	753	2	Q8ZS77	ANASP
77	37	66.1	753	2	Q8YW16	ANASP
78	37	66.1	814	2	Q4FRC3	9GAMM
79	37	66.1	896	2	Q8TWR9	METAC
80	37	66.1	948	2	Q26312	STRPU
81	37	66.1	948	2	Q8KQL9	SACER
82	37	66.1	973	2	Q6N989	RHOPA
83	37	66.1	1067	1	TR11	STRCO
84	37	66.1	1258	2	Q6GUE5	9GAMM
85	37	66.1	1752	2	Q07265	STRPU
86	37	66.1	2936	2	Q9NKP7	LEIMA
87	36	64.3	67	2	Q9PED6	XYLFA
88	36	64.3	96	2	Q7TM46	9ZZZZ
89	36	64.3	115	2	Q61XC2	CABER
90	36	64.3	141	2	Q700Q3	PSFPU
91	36	64.3	160	2	P74976	XYLFA
92	36	64.3	166	2	Q5T4D2	HUMAN
93	36	64.3	194	1	LYG	EPICO
94	36	64.3	217	2	Q7W3Z4	BORPA
95	36	64.3	238	2	Q9KYB0	STRCO
96	36	64.3	248	2	Q7VUW1	BORPE
97	36	64.3	253	2	Q7P500	FUSNV
98	36	64.3	253	2	Q8RH63	FUSNN
99	36	64.3	255	2	Q7WFC6	BORBR
100	36	64.3	263	1	RPP	XYLFT
101	36	64.3	263	2	Q4FGL4	XYLFA
102	36	64.3	275	2	Q96EZ5	HUMAN
103	36	64.3	282	2	Q5NDF6	HYDEC
104	36	64.3	282	2	Q4FGL2	XYLFA

Q9L144	streptomyc
Q74sa4	versinia pe
Q66hy4	brachydamio
Q668c3	versinia ps
Q7x541	actinoplane
Q869y0	dictyosteli
Q553y8	dictyosteli
Q8ue69	agrobacteri
Q4luso	burkholderi
Q606w1	methylococc
Q9hrc7	halobacteri
P10488	klebsiella
Q67js4	symbiobacte
Q5uq11	minivirus
Q4upy2	xanthomonas
Q4npes	anaeromyxob
Q8p4d1	xanthomonas
Q4h433	solanum tub
Q5on02	entamoeba h
Q5quk4	idiomarina
Q8ceb1	mus musculus
Q92wm5	rhizobium m
Q7wlo2	bordeletella
Q6ze25	oryza sativ
Q8rtw0	uncultured
Q6sf93	uncultured
Q6dfsi	xenopus tro
Q6gp74	xenopus lae
P55772	mus musculus
Q921q6	mus musculus
Q544u5	mus musculus
Q6azw0	brachydamio
P97687	rattus norv
Q8cdv7	mus musculus
Q65j87	bacillus li
Q8nqy9	corynebacte
Q82xb5	streptomyc
Q8yc19	anaebana sp
Q5lpf8	silicibacte
Q51tu2	magnaporthe
Q5smq4	oryza sativ
P37279	synchococc
Q8kpv5	synchococc
Q5n650	synchococc
Q8z577	anaebana sp
Q8yw16	anaebana sp
Q4frc3	psychrobact
Q8twn9	methanosa
Q26312	strongyloce
Q8kql9	saccharopol
Q6n989	rhodopseuo
Q6ndez	streptomyc
Q9nkp7	leishmania
Q7tm46	uncultured
Q61xc2	caenorhabdi
Q700q3	pseudomonas
P74976	xyella fas
Q5t4d2	homo sapien
Q90x99	epinephelus
Q7w3z4	bordeletella
Q9kyb0	streptomyc
Q7vuw1	bordeletella
Q7p500	fusobacteri
Q8rh63	fusobacteri
Q7wfc6	bordeletella
Q83wl3	xyella fas
Q4fgl4	xyella fas
Q96ez5	homo sapien
Q5ndf6	hydractinia
Q4fgl2	xyella fas

105	36	64.3	283	1	RPP_XYLFA	Q9phk6	xylella fas	178	35	62.5	252	2	Q7VTS9	BORPE	Q7vts9	bordetella
106	36	64.3	283	2	Q61X02_CABBR	Q6ix02	caenorhabdi	179	35	62.5	257	2	Q8YUW3	ANASP	Q8yuw3	anabaena sp
107	36	64.3	283	2	Q4ZZA7_PSESY	Q4zza7	pseudomonas	180	35	62.5	261	2	Q6ZGL0	ORYSA	Q6zgl0	oryza sativ
108	36	64.3	318	2	Q4UNX9_XANCP	Q4unx9	xanthomonas	181	35	62.5	264	2	Q4ZVM0	PSESY	Q4zvm0	pseudomonas
109	36	64.3	318	2	Q8P3P7_XANCP	Q8p3f7	xanthomonas	182	35	62.5	264	2	Q4IWX7_AZOVI	Q4iwx7	azotobacter	
110	36	64.3	322	2	Q7R524_GIALA	Q7r524	giardia lam	183	35	62.5	264	2	Q8I7S1_PSESM	Q8i7s1	pseudomonas	
111	36	64.3	326	2	Q4K9E6_PSEF5	Q4k9e6	pseudomonas	184	35	62.5	265	2	Q89V33_BRAJA	Q89v33	bradyrhizob	
112	36	64.3	336	2	Q9RDM2_STRCO	Q9rdm2	streptomyces	185	35	62.5	273	2	Q4MYU9_THEPA	Q4myu9	theileria p	
113	36	64.3	336	2	Q82B26_STRAW	Q82bz6	streptomyces	186	35	62.5	276	2	Q6Z378_ORYSA	Q6z378	oryza sativ	
114	36	64.3	338	2	Q9RTH5_DEIRA	Q9rth5	deinococcus	187	35	62.5	283	2	Q87V57_PSESM	Q87v57	pseudomonas	
115	36	64.3	355	2	Q882A5_PSESM	Q882a5	pseudomonas	188	35	62.5	290	2	Q67QM5_SYMBH	Q67qm5	ymbiobacte	
116	36	64.3	356	2	Q912R6_PSEAE	Q912r6	pseudomonas	189	35	62.5	294	2	Q92N10_RHIME	Q92n10	rhizobium m	
117	36	64.3	365	2	Q4N12R6_PSEAE	Q4n12r6	pseudomonas	190	35	62.5	300	2	P74651_SYNY3	P74651	synecocyst	
118	36	64.3	375	2	Q61BM4_CABBR	Q61bm4	caenorhabdi	191	35	62.5	302	2	Q9CBP4_MYCLE	Q9cbp4	mycobacteri	
119	36	64.3	390	2	Q7PUJ5_ANOGA	Q7puj5	anopheles g	192	35	62.5	309	2	Q5WBZ3_BACSK	Q5wbz3	bacillus cl	
120	36	64.3	391	2	Q8WR22_ANOGA	Q8wr22	anopheles g	193	35	62.5	312	2	Q829K6_STRAW	Q829k6	streptomyces	
121	36	64.3	415	2	Q7TZO1_MYCBO	Q7tzo1	mycobacteri	194	35	62.5	319	2	Q8MVC6_IXOSC	Q8mvc6	ixodes scap	
122	36	64.3	415	2	P71981_MYCTU	P71981	mycobacteri	195	35	62.5	332	2	Q69231_BACI	Q69231	bacillus sp	
123	36	64.3	415	2	Q85X61_CRYNE	Q85x61	cryptococcus	196	35	62.5	335	2	Q4JVV1_CORJK	Q4jvv1	corynebacte	
124	36	64.3	424	2	Q5XK61_CRYNE	Q5xk61	cryptococcus	197	35	62.5	341	2	Q9ZET3_XANP2	Q9zet3	xanthobacte	
125	36	64.3	424	2	Q5XK61_CRYNE	Q5xk61	cryptococcus	198	35	62.5	354	2	Q5LNT0_SILPO	Q5lnt0	silicibacte	
126	36	64.3	436	2	Q5F2N5_RAT	Q5f2n5	rattus norv	199	35	62.5	360	2	Q7URK3_RHOBA	Q7urk3	rhodopirell	
127	36	64.3	462	2	Q94BM9_ARATH	Q94bm9	arabidopsis	200	35	62.5	361	2	Q6LHL8_PHOPR	Q6lhl8	photobacter	
128	36	64.3	476	2	Q7SD10_NEUCR	Q7sd10	neurospora	201	35	62.5	364	1	HEMH_CHRVO	HEMH	chromobacte	
129	36	64.3	479	2	Q8N1ML6_ARATH	Q8n1ml6	arabidopsis	202	35	62.5	370	2	Q9HYJ0_PSEAE	Q9hyj0	pseudomonas	
130	36	64.3	517	2	Q8N107_HUMAN	Q8n107	homo sapien	203	35	62.5	379	2	Q4J471_AZOVI	Q4j471	azotobacter	
131	36	64.3	536	2	Q93JK8_STRCO	Q93jkr8	streptomyces	204	35	62.5	383	2	Q6I609_ORYSA	Q6i609	oryza sativ	
132	36	64.3	540	2	Q6C246_YARLI	Q6c246	yarrowia li	205	35	62.5	387	2	Q98FX1_RHILO	Q98fx1	rhizobium l	
133	36	64.3	560	2	Q9P226_HUMAN	Q9p226	homo sapien	206	35	62.5	396	2	Q943M0_ORYSA	Q943m0	oryza sativ	
134	36	64.3	593	2	Q4KMP2_HUMAN	Q4kmp2	homo sapien	207	35	62.5	405	2	Q6RK21_ERWCT	Q6rk21	erwinia car	
135	36	64.3	593	2	Q5T4B2_HUMAN	Q5t4b2	homo sapien	208	35	62.5	412	2	Q9VZV5_DROME	Q9vzv5	drosophila	
136	36	64.3	629	2	Q4IQB8_GIBZE	Q4iqb8	gibberella	209	35	62.5	426	2	Q4ZRV2_PSESY	Q4zrv2	pseudomonas	
137	36	64.3	663	2	Q7SCL7_NEUCR	Q7scl7	neurospora	210	35	62.5	426	2	Q880G9_PSESM	Q880g9	pseudomonas	
138	36	64.3	673	2	Q6VT14_9VIRU	Q6vt14	vibrio para	211	35	62.5	427	1	CYB6_CHLLT	CYB6	chlorobium	
139	36	64.3	675	2	Q6VT78_9VIRU	Q6vt78	vibrio para	212	35	62.5	429	2	Q6Z370_ORYSA	Q6z370	oryza sativ	
140	36	64.3	688	2	Q4II41_GIBZE	Q4ii41	gibberella	213	35	62.5	437	2	Q765A8_9PEZI	Q765a8	gibberella	
141	36	64.3	694	2	Q4NVS9_9DELT	Q4nvs9	anaeromyxob	214	35	62.5	468	2	Q4I330_GIBZE	Q4i330	gibberella	
142	36	64.3	701	1	ACOX3_YARLI	Q74936	yarrowia li	215	35	62.5	480	2	Q8KF93_CHLTE	Q8kf93	chlorobium	
143	36	64.3	701	2	Y4II_RHISN	P55492	rhizobium s	216	35	62.5	482	2	Q04201_ARATH	Q04201	arabidopsis	
144	36	64.3	703	1	Q5YVJ8_NOCFA	P55492	rhizobium s	217	35	62.5	495	2	Q9H9A9_HUMAN	Q9h9a9	homo sapien	
145	36	64.3	742	2	Q5YVJ8_NOCFA	P55492	rhizobium s	218	35	62.5	495	2	Q84CJ1_9ACTO	Q84cj1	streptomyces	
146	36	64.3	770	2	Q4IYM9_AZOVI	Q4iym9	azotobacter	219	35	62.5	520	2	Q9H0W8_HUMAN	Q9h0w8	homo sapien	
147	36	64.3	951	2	Q8HTCA_PSEAE	Q8htc4	pseudomonas	220	35	62.5	520	2	Q988D6_RHILO	Q988d6	rhizobium l	
148	36	64.3	1260	1	LINI_NYCCO	P08548	nycticebus	221	35	62.5	520	2	Q5PQ86_RAT	Q5pq86	rattus norv	
149	35	62.5	44	2	Q4P6D7_USTMA	Q4p6d7	trypanosoma	222	35	62.5	520	2	Q8BYJ3_MOUSE	Q8byj3	mus musculu	
150	35	62.5	66	2	Q4GY73_9TRYP	Q4gy73	trypanosoma	223	35	62.5	520	2	Q9DB90_MOUSE	Q9db90	mus musculu	
151	35	62.5	86	2	Q88GUL_PSEPK	Q88gul	pseudomonas	224	35	62.5	527	2	Q4ZTC1_PSESY	Q4ztc1	pseudomonas	
152	35	62.5	91	2	Q7QMC5_GIALA	Q7qwc5	giardia lam	225	35	62.5	534	2	Q4SVX6_TETNG	Q4svx6	tetradodon n	
153	35	62.5	91	2	Q5GR82_ALCXX	Q5gr82	alcaligenes	226	35	62.5	542	2	Q55KX2_CRYNE	Q55kx2	cryptococcus	
154	35	62.5	113	2	Q9KT37_VIBCH	Q9kt37	vibrio chol	227	35	62.5	542	2	Q5KAK7_CRYNE	Q5kak7	cryptococcus	
155	35	62.5	117	2	Q7NKS7_CHRVO	Q7nks7	chromobacte	228	35	62.5	545	2	Q6CAR7_YARLI	Q6car7	yarrowia li	
156	35	62.5	130	2	Q27408_METHV	Q27408	methanobact	229	35	62.5	547	2	Q55KX1_CRYNE	Q55kx1	cryptococcus	
157	35	62.5	130	2	Q58KA3_HCMV	Q58ka3	human cytom	230	35	62.5	547	2	Q5KAK8_CRYNE	Q5kak8	cryptococcus	
158	35	62.5	130	2	Q8QH04_HCMV	Q8qh04	human cytom	231	35	62.5	553	2	Q55YK8_CRYNE	Q55yk8	cryptococcus	
159	35	62.5	130	2	Q8QH05_HCMV	Q8qh05	human cytom	232	35	62.5	553	2	Q5KLD0_CRYNE	Q5kld0	cryptococcus	
160	35	62.5	130	2	Q8QR26_HCMV	Q8qr26	human cytom	233	35	62.5	557	2	Q55NQ3_CRYNE	Q55nq3	cryptococcus	
161	35	62.5	130	2	Q6SW19_HCMV	Q6sw19	human cytom	234	35	62.5	557	2	Q5KAC3_CRYNE	Q5kac3	cryptococcus	
162	35	62.5	130	2	Q6SWL0_HCMV	Q6swl0	human cytom	235	35	62.5	586	2	Q4LX14_9BURK	Q4lx14	burkholderi	
163	35	62.5	130	2	Q6SWU3_HCMV	Q6swu3	human cytom	236	35	62.5	607	1	GLTLA_HUMAN	GLTLA	homo sapien	
164	35	62.5	130	2	Q6SWU3_HCMV	Q6swu3	human cytom	237	35	62.5	607	2	Q58A54_HUMAN	Q58a54	homo sapien	
165	35	62.5	158	2	Q7VX29_BORPE	Q7vx29	bordetella	238	35	62.5	607	2	Q59J91_MOUSE	Q59j91	mus musculu	
166	35	62.5	158	2	Q7WLYS_BORPA	Q7wly5	bordetella	239	35	62.5	622	1	GLTLA_MOUSE	GLTLA	m putative	
167	35	62.5	158	2	Q7WLYS_BORBR	Q7wly5	bordetella	240	35	62.5	641	2	Q63JQ4_BURPS	Q63jq4	burkholderi	
168	35	62.5	195	2	Q8U3H5_PYRFP	Q8u3h5	pyrococcus	241	35	62.5	641	2	Q62AN3_BURMA	Q62an3	burkholderi	
169	35	62.5	213	2	Q88JD7_PSEPK	Q88jd7	pseudomonas	242	35	62.5	647	2	Q82L15_STRAW	Q82l15	streptomyces	
170	35	62.5	226	2	Q82QP7_STRAM	Q82qp7	streptomyces	243	35	62.5	697	1	EPG2_SHEON	EPG2	shewanella	
171	35	62.5	227	2	Q836F0_ENTFA	Q836f0	enterococcus	244	35	62.5	702	2	Q52ET5_MAGGR	Q52et5	magnaporthe	
172	35	62.5	228	2	Q5TN19_ANOGA	Q5tni9	anopheles g	245	35	62.5	702	2	Q7UY61_RHOBA	Q7uy61	rhodopirell	
173	35	62.5	233	2	Q6N7U1_RHOBA	Q6n7u1	rhodopsin	246	35	62.5	723	2	Q5IND2_MAGGR	Q5ind2	magnaporthe	
174	35	62.5	238	2	Q5ZAL0_ORYSA	Q5zal0	oryza sativ	247	35	62.5	723	2	Q4UXI0_XANCP	Q4uxi0	xanthomonas	
175	35	62.5	241	2	Q4NRG0_9DELT	Q4nrg0	anaeromyxob	248	35	62.5	723	2	Q8P6N1_XANCP	Q8p6n1	xanthomonas	
176	35	62.5	251	2	Q6IU32_CABBR	Q6iu32	caenorhabdi	249	35	62.5	731	2	Q8PHY5_XANAC	Q8phy5	xanthomonas	
177	35	62.5	252	2	Q7WQD4_BORBR	Q7wqd4	bordetella	250	35	62.5	731	2	Q6N6R0_RHOBA	Q6n6r0	rhodopsin	

251	35	62.5	733	2	Q92ZA2_RHIME	Q92za2 rhizobium m	324	34	60.7	282	2	Q55VX0_CRYNE	Q55vx0 cryptococcus
252	35	62.5	740	2	Q9RJD7_STRCO	Q9rjd7 streptomyce	325	34	60.7	282	2	Q5KKC0_CRYNE	Q5kkc0 cryptococcus
253	35	62.5	749	2	Q825G4_STRAW	Q825g4 streptomyce	326	34	60.7	297	2	Q7SG71_NEUCR	Q7sg71 neurospora
254	35	62.5	755	2	Q5SR90_CRYNE	Q5sr90 cryptococcus	327	34	60.7	305	2	Q4MDL1_RHOPA	Q4mdl1 rhodopseuod n
255	35	62.5	755	2	Q5KEY0_CRYNE	Q5key0 cryptococcus	328	34	60.7	308	2	Q4MR1_TETNG	Q4mr1 tetraodon n
256	35	62.5	765	2	Q5H231_XANOR	Q5h231 xanthomonas	329	34	60.7	311	2	Q577B0_BRUAB	Q577b0 brucella ab
257	35	62.5	771	2	Q4P3Q5_USTMA	Q4p3q5 ustilago ma	330	34	60.7	313	2	Q6FT29_CANGA	Q6ft29 candida gla
258	35	62.5	787	2	Q74BT7_GEOSL	Q74bt7 geobacter s	331	34	60.7	314	2	Q61FB9_CABBR	Q61fb9 caenorhabdi
259	35	62.5	824	2	Q67PT7_SYNTH	Q67pt7 symbiobacte	332	34	60.7	316	2	Q72756_HUMAN	Q72756 homo sapien
260	35	62.5	844	2	Q9V9G5_DROME	Q9v9g5 drosophila	333	34	60.7	337	2	Q20269_CABEL	Q20269 caenorhabdi
261	35	62.5	891	2	Q9RWY5_DEIRA	Q9rwy5 deinococcus	334	34	60.7	337	2	Q7UT17_RHOBA	Q7ut17 rhodopirell
262	35	62.5	1051	2	P97961_9PEZI	P97961 cylindrotriti	335	34	60.7	340	2	Q62187_BURMA	Q62187 burkholderi
263	35	62.5	1054	2	Q726E5_HUMAN	Q726e5 homo sapien	336	34	60.7	340	2	Q63V93_BURPS	Q63v93 burkholderi
264	35	62.5	1100	2	Q5CBE3_9THEM	Q5cbe3 thermotoga	337	34	60.7	343	2	Q84555_CHLTR	Q84555 chlamydia t
265	35	62.5	1100	2	Q5CBG9_9THEM	Q5cbg9 thermotoga	338	34	60.7	345	2	Q8SQX7_ENCPU	Q8sqx7 encephalito
266	35	62.5	1100	2	Q5CBS3_9THEM	Q5cbs3 thermotoga	339	34	60.7	349	2	Q84ZS1_ORYSA	Q84zsl oryza sativ
267	35	62.5	1103	2	Q5CBC0_9THEM	Q5cbc0 thermotoga	340	34	60.7	355	2	Q7S447_NEUCR	Q7s447 neurospora
268	35	62.5	1104	1	RGYR_THEMEA	Q51934 thermotoga	341	34	60.7	367	2	Q806B3_CHV1	Q806b3 cercopithe
269	35	62.5	1104	2	Q5CB68_9THEM	Q5cb68 thermotoga	342	34	60.7	375	2	Q7TSB6_CHV1	Q7tsb6 cercopithe
270	35	62.5	1104	2	Q5CBJ6_9THEM	Q5cbj6 thermotoga	343	34	60.7	380	2	Q5YTL0_NOCFA	Q5ytl0 nocardia fa
271	35	62.5	1104	2	Q5CBL4_9THEM	Q5cbl4 thermotoga	344	34	60.7	380	2	Q5P902_AZOSE	Q5p902 azoarcus sp
272	35	62.5	1117	1	RGYR_THETN	Q5cb14 thermotoga	345	34	60.7	387	2	Q5LPN8_SILPO	Q5lpn8 silicibacte
273	35	62.5	1119	2	Q5CBQ4_THENE	Q5cbq4 thermotoga	346	34	60.7	388	2	Q6V1M5_9ACTO	Q6v1m5 streptomyce
274	35	62.5	1122	2	Q5PAV7_AZOSE	Q5pav7 azoarcus sp	347	34	60.7	388	2	Q7U9B2_SYNFX	Q7u9b2 synchococc
275	35	62.5	1157	2	Q51HR4_MAGGR	Q51hr4 magnaporthe	348	34	60.7	400	1	CSD_CHIMU	Q9plp0 chlamydia m
276	35	62.5	1223	2	Q4PFA4_USTMA	Q4pfa4 ustilago ma	349	34	60.7	402	2	Q86GV0_CABEL	Q86gv0 caenorhabdi
277	35	62.5	1226	2	Q69Z42_MOUSE	Q69z42 mus musculus	351	34	60.7	407	2	Q4J1R9_AZOVI	Q4j1r9 azotobacter
278	35	62.5	1227	2	Q7MWZ3_PHOLL	Q7mwz3 photorhabdu	352	34	60.7	408	2	Q631F3_BURPS	Q631f3 burkholderi
279	35	62.5	1238	2	Q5V159_MOUSE	Q5v159 mus musculus	353	34	60.7	415	2	Q62DP9_BURMA	Q62dp9 burkholderi
280	35	62.5	1273	2	Q9YCAL_AERPE	Q9ycal aeropyrum p	354	34	60.7	415	2	Q7VXC4_BORPE	Q7vxc4 bordetella
281	35	62.5	1494	2	Q4LVA6_BHOBK	Q4lva6 burkholderi	355	34	60.7	415	2	Q7WA23_BORPA	Q7wa23 bordetella
282	35	62.5	1631	2	Q7UL61_RHOBK	Q7ul61 rhodopirell	356	34	60.7	430	2	Q7WJ52_BORBR	Q7wj52 bordetella
283	35	62.5	3007	2	Q63S02_BURPS	Q63su2 burkholderi	357	34	60.7	434	2	Q5NWB8_AZOSE	Q5nwb8 azoarcus sp
284	35	62.5	4306	2	Q9UJ79_RAT	Q9uj79 rattus norv	358	34	60.7	437	2	Q6NBO6_RHOPA	Q6nbq6 rhodopseuod
285	34.5	61.6	151	2	Q93WV9_MUSAC	Q93wv9 musa acumin	359	34	60.7	449	2	Q5TWX9_ANOGA	Q5twx9 anopheles g
286	34.5	61.6	341	1	QHEA_CLOPE	Q8xj15 clostridium	360	34	60.7	457	1	ARDY_ECO57	Q8x730 escherichia
287	34.5	61.6	433	2	Q4H653_9DEIO	Q4h653 deinococcus	361	34	60.7	457	1	ARDY_ECOL6	Q8fb96 escherichia
288	34.5	61.6	550	2	Q5SV18_CRYNE	Q5sv18 cryptococcus	362	34	60.7	457	1	ARDY_ECOL6	P11447 escherichia
289	34.5	61.6	550	2	Q5KKQ1_CRYNE	Q5kkq1 cryptococcus	363	34	60.7	457	1	ARDY_SHIFL	P59619 shigella fl
290	34.5	61.6	671	2	Q47867_PANAY	Q47867 pantoea agg	364	34	60.7	457	1	ARDY_YERPE	Q8za88 versinia pe
291	34.5	61.6	671	2	Q85671_ENTAG	Q85671 pantoea agg	365	34	60.7	457	1	ARDY_YERPE	Q66g70 versinia ps
292	34.5	61.6	1032	2	Q74CR1_GEOSL	Q74cr1 geobacter s	366	34	60.7	458	1	ARDY_SALTY	Q8zkl6 salmonella
293	34	60.7	76	2	Q9BM17_RHIL0	Q9bm17 rhizobium l	367	34	60.7	458	1	ARDY_SALTY	Q8zkl6 salmonella
294	34	60.7	145	2	Q5GWE1_XANOR	Q5gwe1 xanthomonas	368	34	60.7	458	2	Q57H93_SALCH	Q57h93 salmonella
295	34	60.7	146	2	Q877E8_9CREN	Q877e8 hyperthermu	369	34	60.7	458	2	Q5PK73_SALPA	Q5pk73 salmonella
296	34	60.7	146	2	Q6L3Z1_SOLDE	Q6l3z1 solanum dem	370	34	60.7	471	2	Q82NF9_SITAW	Q82nf9 streptomyce
297	34	60.7	152	2	Q4UR61_XANCP	Q4ur61 xanthomonas	371	34	60.7	475	2	Q7MR32_WOLSU	Q7mr32 wolinnella s
298	34	60.7	152	2	Q8PCC2_XANCP	Q8pcc2 xanthomonas	372	34	60.7	477	2	Q01870_CABEL	Q01870 caenorhabdi
299	34	60.7	154	1	RNH_BRUME	P66673 brucella me	373	34	60.7	488	2	O87193_HELPY	O87193 helicobacte
300	34	60.7	154	1	RNH_BRUSU	P66674 brucella su	374	34	60.7	488	2	O92MT6_HELPJ	O92mt6 helicobacte
301	34	60.7	154	2	Q57EP4_BRUAB	Q57ep4 brucella ab	375	34	60.7	490	1	ARDY_BTFL0	O8g5f3 bifidobacte
302	34	60.7	168	2	Q4LZK2_9BURK	Q4lzk2 burkholderi	376	34	60.7	491	2	Q8QW2_PSEPA	Q8qkw2 pseudomonas
303	34	60.7	176	2	Q5V3R0_HALMA	Q5v3r0 haloarcula	377	34	60.7	517	2	Q9F536_ECOLI	Q9f536 escherichia
304	34	60.7	192	2	Q8PP03_XANAC	Q8pp03 xanthomonas	378	34	60.7	518	2	Q72CW0_DESVH	Q72cw0 desulfovibr
305	34	60.7	193	2	Q916T7_PSEAE	Q916t7 pseudomonas	379	34	60.7	524	2	Q8LSK4_PHYPA	Q8lsk4 physocitre
306	34	60.7	198	2	Q6YV27_ORYSA	Q6yv27 oryza sativ	380	34	60.7	526	2	Q5OX16_IDILO	Q5ox16 idiomarina
307	34	60.7	198	2	Q9MAH8_VITVI	Q9mah8 vitis vinif	381	34	60.7	531	1	KPKY_EIMTE	O44006 eimeria ten
308	34	60.7	201	2	Q6ARE4_DESPS	Q6are4 desulfotale	382	34	60.7	544	2	Q7NE97_GLOVI	Q7ne97 gloobacter
309	34	60.7	202	2	Q8NA16_HUMAN	Q8na16 homo sapien	383	34	60.7	548	1	E13B_ARTSW	Q59146 arthrobacte
310	34	60.7	203	1	PDXT_AERPE	Q9yfk4 aeropyrum p	384	34	60.7	548	1	E13B_ARTSW	P22222 oerokovia x
311	34	60.7	213	1	KGUA_CAUCR	Q9a7n9 caulobacter	385	34	60.7	549	1	E13B_ARTSW	Q51vp7 magnaporthe
312	34	60.7	214	2	Q7W955_BORPA	Q7w295 bordetella	386	34	60.7	556	2	Q51VP7_MAGGR	Q4k416 pseudomonas
313	34	60.7	214	2	Q7WR62_BORBR	Q7wr62 bordetella	387	34	60.7	568	2	O4K416_PSEF5	O6u4f4 psittacid h
314	34	60.7	225	1	PGP_HALSA	Q6ab6 propionibac	388	34	60.7	569	1	PEX5_PICAN	Q01495 picchia angu
315	34	60.7	233	2	Q6ABB6_PROAC	Q6ab6 propionibac	389	34	60.7	581	2	Q4P3Z0_USTMA	Q93cy5 thiocopea r
316	34	60.7	235	2	Q7MVZ1_PORGI	Q7mvz1 porphyromon	390	34	60.7	591	2	Q9FNB3_ARATH	Q4p3z0 ustilago ma
317	34	60.7	237	2	Q7ZV80_BRARE	Q7zv80 brachydanio	391	34	60.7	596	2	Q529E7_MAGGR	Q9fnb3 arabidopsis
318	34	60.7	245	2	Q4NUG0_9DELT	Q4nug0 anaeromyxob	392	34	60.7	596	2	Q529E7_MAGGR	Q529e7 magnaporthe
319	34	60.7	247	2	Q4KGE8_PSEF5	Q4kge8 pseudomonas	393	34	60.7	596	2	Q4S7P7_TETNG	Q4s7p7 tetraodon n
320	34	60.7	252	2	Q9ABY3_CAUCR	Q9aby3 caulobacter	394	34	60.7	598	2	Q8H78_PSEPK	Q8h78 pseudomonas
321	34	60.7	260	2	Q6O419_METCA	Q6o419 methylococc	395	34	60.7	613	2	Q5LWS0_SILPO	Q5lws0 silicibacte
322	34	60.7	271	2	Q8YBE6_BRUME	Q8ybe6 brucella me	396	34	60.7				
323	34	60.7	273	2	Q66T06_KLEPN	Q66t06 klebsiella							

397	34	60.7	618	2	Q8W4C3_ARATH	Q8w4c3_arabidopsis	470	33	58.9	154	2	Q8TLD3_METAC	Q8tld3_methanosarc	
398	34	60.7	644	2	Q7XND4_ORYSA	Q7xnd4_oryza sativ	471	33	58.9	154	2	Q4LYN3_9BURK	Q4lyn3_burkholderi	
399	34	60.7	647	2	Q8Z5E9_STRAW	Q8z5e9_streptomyce	472	33	58.9	179	2	Q9HZH7_PSEAE	Q9hzh7_pseudomonas	
400	34	60.7	659	2	Q5ZCG0_ORYSA	Q5zcg0_oryza sativ	473	33	58.9	181	2	Q4RZR7_TETNG	Q4rzz7_tetradon n	
401	34	60.7	660	2	Q17248_BOOMI	Q17248_boophilus m	474	33	58.9	185	2	Q5JDD5_PYRKO	Q5jdd5_pyrococcus	
402	34	60.7	693	2	Q8ZQ94_SALTY	Q8zq94_salmonella	475	33	58.9	187	2	Q629E8_CAEBR	Q629e8_caenorhabdi	
403	34	60.7	703	2	Q4WCY1_ASPFU	Q4wcy1_aspergillus	476	33	58.9	190	2	Q9UY56_PYRAB	Q9uy56_pyrococcus	
404	34	60.7	725	2	Q5XVM9_ENTFA	Q5xvm9_enterococcu	477	33	58.9	196	2	Q7WZ84_9ACTO	Q7wz84_nonomurasa	
405	34	60.7	733	2	Q89LV4_BRAJA	Q89lv4_bradyrhizob	478	33	58.9	196	2	Q5X2K8_LEGPA	Q5x2k8_legionella	
406	34	60.7	738	2	Q8XYC9_RALSO	Q8xyc9_raistonia s	479	33	58.9	197	1	HAW1_CHRYO	Q7nzj6_chromobacte	
407	34	60.7	750	2	Q9FGC2_ARATH	Q9fgc2_arabidopsis	480	33	58.9	199	2	Q9AUC9_SESIN	Q9auc9_sesamum ind	
408	34	60.7	856	2	Q4Q466_LEIMA	Q4q466_leishmania	481	33	58.9	200	2	Q51RL6_MAGGR	Q51rl6_magnaporthe	
409	34	60.7	859	2	Q9FW70_ORYSA	Q9fw70_oryza sativ	482	33	58.9	200	2	Q7FAY5_ORYSA	Q7fay5_oryza sativ	
410	34	60.7	870	2	Q7QJD3_ANOGA	Q7qjd3_anopheles g	483	33	58.9	200	2	Q828G1_STRAW	Q828g1_streptomyce	
411	34	60.7	901	1	MALT_SALTY	Q8z227_salmonella	484	33	58.9	200	2	Q8CJ54_STRCO	Q8cjs4_streptomyce	
412	34	60.7	901	1	MALT_SALTY	Q8z227_salmonella	485	33	58.9	214	2	Q7VT72_BORPE	Q7vt72_bordetella	
413	34	60.7	901	2	Q57IV9_SALCH	Q57iv9_salmonella	486	33	58.9	215	2	Q96Y57_SULTO	Q96y57_sulfolobus	
414	34	60.7	901	2	Q5PM00_SALPA	Q5pm00_salmonella	487	33	58.9	215	2	Q4WH75_ASPFU	Q4wh75_aspergillus	
415	34	60.7	905	2	Q8XRT2_RALSO	Q8xrt2_raistonia s	488	33	58.9	224	2	Q51KN9_MAGGR	Q51kn9_magnaporthe	
416	34	60.7	906	2	Q630Y0_BACZ	Q630y0_bacillus ce	489	33	58.9	225	2	Q9HVR1_PSEAE	Q9hvr1_pseudomonas	
417	34	60.7	987	2	Q7YZM8_CABEL	Q7yzm8_caenorhabdi	490	33	58.9	232	2	Q840V0_SPHEL	Q840v0_sphingomona	
418	34	60.7	1051	2	Q9Y150_DROME	Q9y150_drosophila	491	33	58.9	240	2	Q57891_PYRHO	Q57891_pyrococcus	
419	34	60.7	1058	2	Q4WV56_ASPFU	Q4wv56_aspergillus	492	33	58.9	240	2	Q6CL65_KULUA	Q6cl65_kluyveromyc	
420	34	60.7	1076	2	Q605X2_METCA	Q605x2_methyllococc	493	33	58.9	243	2	Q7NN94_GLOVI	Q7nn94_gloeobacter	
421	34	60.7	1098	2	Q7Z146_CABEL	Q7z146_caenorhabdi	494	33	58.9	244	2	Q9RWM1_DEIRA	Q9rwm1_deinococcus	
422	34	60.7	1108	2	Q4X0Z7_ASPFU	Q4x0z7_aspergillus	495	33	58.9	248	2	Q5P7A3_AZOSE	Q5p7a3_azococcus sp	
423	34	60.7	1112	2	Q8GD29_BORAV	Q8gd29_bordetella	496	33	58.9	249	2	Q8X6D3_ECO57	Q8x6d3_escherichia	
424	34	60.7	1115	2	Q9LSV7_ARATH	Q9lsv7_arabidopsis	497	33	58.9	257	2	Q5LU90_SILPO	Q5lu90_silicibacte	
425	34	60.7	1116	2	Q53VZ7_THET8	Q53vz7_thermus the	498	33	58.9	258	2	Q4PMJ6_IXOSC	Q4pmj6_ixodes scap	
426	34	60.7	1159	1	RGYR2_AQUAE	Q67226_aquifex aeo	499	33	58.9	261	2	Q87AQ1_XYLEF	Q87aq1_xylella fas	
427	34	60.7	1166	1	RGYR2_SULSO	Q7z4f5_sulfolobus	500	33	58.9	261	2	Q9PEW5_XYLEFA	Q9pew5_xylella fas	
428	34	60.7	1166	1	RGYR_SULSH	P74759_sulfolobus	501	33	58.9	261	2	Q7X4S4_BACLI	Q7x4s4_bacillus li	
429	34	60.7	1192	2	Q610E4_CAEBR	Q610e4_caenorhabdi	502	33	58.9	265	2	Q65PM6_BACLD	Q65fm6_bacillus li	
430	34	60.7	1208	2	Q7RMM1_NEUCR	Q7rmm1_neurospora	503	33	58.9	265	2	Q8KJ89_RHILO	Q8kj89_rhizobium l	
431	34	60.7	1214	1	RGYR_PYRAB	Q9uz86_pyrococcus	504	33	58.9	272	2	Q7TW42_MYCBO	Q7tw42_mycobacteri	
432	34	60.7	1214	1	RGYR_PYRFU	P95479_pyrococcus	505	33	58.9	272	2	Q06282_MYCTO	Q06282_mycobacteri	
433	34	60.7	1214	1	Q49600_METKA	Q49600_methanopyru	506	33	58.9	273	2	Q5P6P9_AZOSE	Q5p6p9_azococcus sp	
434	34	60.7	1221	2	Q49600_METKA	Q49600_methanopyru	507	33	58.9	275	2	Q9A3U3_CAUCR	Q9a3u3_caulobacter	
435	34	60.7	1222	1	RGYR2_AERPE	Q9yc75_aeropyrum p	508	33	58.9	281	2	Q63X73_BURPS	Q63x73_burkholderi	
436	34	60.7	1228	1	RGYR_PYRAB	Q8xt5_pyrobaculum	509	33	58.9	282	1	FFG_PROMA	Q7vdk6_prochloroco	
437	34	60.7	1332	1	YSY1_CABEL	Q20255_caenorhabdi	510	33	58.9	287	2	Q7TXN3_MYCBO	Q7txn3_mycobacteri	
438	34	60.7	1379	2	Q581S3_9ACTO	Q581s3_streptomyce	511	33	58.9	287	2	Q7D6F5_MYCTO	Q7d6f5_mycobacteri	
439	34	60.7	1529	2	Q9Y7C6_ASPTE	Q9y7c6_aspergillus	512	33	58.9	288	2	Q5PBF3_ANAMM	Q5pbf3_anaplasma m	
440	34	60.7	1613	1	RGYR_METJA	Q58907_methanococc	513	33	58.9	288	2	Q8Y216_RALSO	Q8y216_ralstonia s	
441	34	60.7	1624	1	RGYR_PYRHO	Q58530_pyrococcus	514	33	58.9	290	2	Q5YV19_NOCFA	Q5yvi9_nocardia fa	
442	34	60.7	1711	1	RGYR_PYRKO	Q6f598_pyrococcus	515	33	58.9	292	2	Q6ZML9_BURMA	Q6zml9_burkholderi	
443	34	60.7	1897	2	Q5L2K6_GEOKA	Q5l2k6_geobacillus	516	33	58.9	293	2	Q8P114_XANAC	Q8p114_xanthomonas	
444	34	60.7	2066	2	Q923T8_PSESX	Q923t8_pseudomonas	517	33	58.9	295	2	Q4LU32_9BURK	Q4lu32_burkholderi	
445	34	60.7	2066	2	Q8W69_PSESM	Q8w69_pseudomonas	518	33	58.9	299	2	Q4TMS3_9SPHN	Q4tms3_erythrobact	
446	34	60.7	2090	2	Q9P6X4_NEUCR	Q9p6x4_neurospora	519	33	58.9	302	2	Q7NV73_CHRVO	Q7nv73_chromobacte	
447	34	60.7	2219	2	Q4P405_USTMA	Q4p405_ustilago ma	520	33	58.9	305	1	RNZ_METMA	Q8q032_methanosarc	
448	33.5	59.8	448	152	2	Q8GGP3_STRAZ	Q8gpp3_streptomyce	521	33	58.9	310	2	Q5NNL6_ZYMMO	Q5nnl6_zymomonas m
449	33.5	59.8	209	2	Q4KPK2_9PSED	Q4kpk2_pseudomonas	522	33	58.9	310	2	Q69LV5_ORYSA	Q69lv5_oryza sativ	
450	33.5	59.8	418	2	Q65H11_BACLD	Q65h11_bacillus li	523	33	58.9	316	2	Q34689_BACSU	Q34689_bacillus su	
451	33.5	59.8	640	2	Q98R10_9RODE	Q98r10_microtus br	524	33	58.9	318	2	Q742P0_MYCPA	Q742p0_mycobacteri	
452	33.5	59.8	688	2	Q5TVL2_ANOGA	Q5tv22_anopheles g	525	33	58.9	320	2	Q4PMW5_IXOSC	Q4pmw5_ixodes scap	
453	33	58.9	38	2	Q6V7N1_9CAUD	Q6v7n1_burkholderi	526	33	58.9	327	2	Q4WC85_ASPFU	Q4wc85_aspergillus	
454	33	58.9	38	2	Q8KC11_CHLTE	Q8kc11_chlorobium	527	33	58.9	327	2	Q41IG2_GIBZE	Q41ig2_gibberella	
455	33	58.9	68	2	Q8UKE3_AGR75	Q8uke3_agrobacteri	528	33	58.9	330	2	Q51L10_MAGGR	Q51l10_magnaporthe	
456	33	58.9	90	2	Q4IU87_AZOV1	Q4iu87_azotobacter	529	33	58.9	330	2	Q4TQX1_9SPHN	Q4tqx1_erythrobact	
457	33	58.9	104	2	Q72CX5_DESVH	Q72cx5_desulfovibr	530	33	58.9	331	2	Q4R3B8_MACFA	Q4r3b8_macaca fasc	
458	33	58.9	112	2	Q6N0E5_RHOPA	Q6n0e5_rhodospheudo	531	33	58.9	333	2	Q8RJ10_STRCO	Q8rj10_streptomyce	
459	33	58.9	112	2	Q4R9Q6_TETNG	Q4r9q6_tetradon n	532	33	58.9	333	2	Q88LW1_PSEPK	Q88lw1_pseudomonas	
460	33	58.9	114	2	Q6W5R6_TRYCR	Q6w5r6_trypanosoma	533	33	58.9	335	2	Q7SHA6_NEUCR	Q7sha6_neurospora	
461	33	58.9	114	2	Q7W775_BORPA	Q7w775_bordetella	534	33	58.9	335	2	Q870E8_9EURO	Q870e8_penicillium	
462	33	58.9	114	2	Q7WIJ7_BORBR	Q7wif7_bordetella	535	33	58.9	335	2	Q9ZWF3_LYCES	Q9zwf3_lycoperisico	
463	33	58.9	125	2	Q5KFE3_CRYNE	Q5kfe3_cryptococcu	536	33	58.9	342	2	Q9X6D4_CLOBE	Q9x6d4_clostridium	
464	33	58.9	130	2	Q8VMF1_PSEPU	Q8vmf1_pseudomonas	537	33	58.9	354	2	Q8PZ82_METMA	Q8pzb2_methanosarc	
465	33	58.9	140	2	Q7P018_CHRVO	Q7p018_chromobacte	538	33	58.9	355	2	Q4ZTM6_PSESY	Q4ztm6_pseudomonas	
466	33	58.9	142	2	Q6N7J4_RHOPA	Q6n7j4_rhodospheudo	539	33	58.9	357	2	Q6LXV5_METMP	Q6lxy5_methanococc	
467	33	58.9	146	2	Q9Y9H4_AERPYR	Q9y9h4_aeropyrum p	540	33	58.9	357	2	Q9A4W7_CAUCR	Q9a4w7_caulobacter	
468	33	58.9	148	2	Q61IP0_DROME	Q61ip0_drosophila	541	33	58.9	359	2	Q852E6_ORYSA	Q852e6_oryza sativ	
469	33	58.9	149	2	Q55QS1_CRYNE	Q55qs1_cryptococcu	542	33	58.9	359	2	Q9HR47_HALSA	Q9hr47_halobacteri	
470	33	58.9	151	2	Q67MR0_SYNTH	Q67mr0_symbiobacte	543	33	58.9	360	2			

543	33	58.9	362	2	Q6J657_ORYSA	Q6J657_oryza sativ	616	33	58.9	570	2	Q7S2U7_NEUCR	Q7a2u7 neurospora
544	33	58.9	362	2	Q5Y0T7_9ALPH	Q5Y0T7 corynebacte	617	33	58.9	573	1	ML011_ARATH	Q9f100 arabidopsis
545	33	58.9	363	2	Q8NQB0_CORGL	Q8NQB0 corynebacte	618	33	58.9	573	1	Q4W8R1_WHEAT	Q4w8r1 triticum ae
546	33	58.9	364	2	Q9R701_9RHIZ	Q9R701 agrobacteri	619	33	58.9	589	2	Q4W8R0_WHEAT	Q4w8r0 triticum ae
547	33	58.9	364	2	Q6P1P5_BURPS	Q6P1P5 burkholderi	620	33	58.9	601	2	Q4H6L3_9DEIO	Q4h6l3 deinococcus
548	33	58.9	365	2	Q4P6I7_USTMA	Q4P6I7 uscillego ma	621	33	58.9	603	2	Q8FPW3_CORFP	Q8fpw3 corynebacte
549	33	58.9	367	2	Q74GD2_GEOSL	Q74GD2 geobacter s	622	33	58.9	603	2	Q9L2I7_STRCO	Q9l2i7 streptomyce
550	33	58.9	371	2	Q98IX0_RHILO	Q98IX0 rhizobium l	623	33	58.9	604	2	Q5B4E5_EBENI	Q5b4e5 aspergillus
551	33	58.9	376	2	Q6VX17_ORYSA	Q6VX17 oryza sativ	624	33	58.9	605	2	Q4NTT6_9DELT	Q4ntt6 anaeromyxob
552	33	58.9	377	2	Q6AFEA_LEIXX	Q6AFEA leifsonia x	625	33	58.9	608	2	Q55WN2_CRYNE	Q55wn2 cryptococcu
553	33	58.9	377	2	Q8XKP0_STRAW	Q8XKP0 streptomyce	626	33	58.9	608	2	Q5KJL3_CRYNE	Q5kjl3 cryptococcu
554	33	58.9	378	2	Q6ASQ2_ORYSA	Q6ASQ2 oryza sativ	627	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
555	33	58.9	384	1	RECF_BRUME	Q6aed7 bruceella me	628	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
556	33	58.9	384	1	RECF_BRUSE	Q8g3e5 bruceella su	629	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
557	33	58.9	384	2	Q7G08_BRUAB	Q8g3e5 bruceella su	630	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
558	33	58.9	389	2	Q7NSI7_WOLSU	Q7nei7 wolinnella s	631	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
559	33	58.9	389	2	Q89NU2_BRADRHIZ	Q8nu2 bradyrhizob	632	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
560	33	58.9	390	2	Q4S165_TETNG	Q4s165 tetradon n	633	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
561	33	58.9	391	2	Q60ER6_ORYSA	Q60er6 oryza sativ	634	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
562	33	58.9	396	1	SOTB_HAEIN	P44535 haemophilus	635	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
563	33	58.9	396	2	Q9HS71_HALSA	Q9hs71 halobacteri	636	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
564	33	58.9	402	2	Q4NNL7_9DELT	Q4nnl7 anaeromyxob	637	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
565	33	58.9	402	2	Q4QP52_HAE18	Q4qp52 haemophilus	638	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
566	33	58.9	404	2	Q5YR35_NOCFA	Q5yr35 nocardia fa	639	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
567	33	58.9	404	2	Q5BDA3_EBENI	Q5bda3 aspergillus	640	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
568	33	58.9	405	2	Q62FRI_BURMA	Q62fri burkholderi	641	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
569	33	58.9	407	2	Q4HU19_GIBZE	Q4hu19 gibberella	642	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
570	33	58.9	421	2	Q89LW8_BRAJA	Q89lw8 bradyrhizob	643	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
571	33	58.9	422	2	Q6NJM3_CORDI	Q6njm3 corynebacte	644	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
572	33	58.9	428	1	CINA_MYCLE	Q5lv46 silicibacte	645	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
573	33	58.9	434	2	Q5LV46_SILPO	Q5lv46 silicibacte	646	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
574	33	58.9	436	2	Q4JU84_CORJK	Q4ju84 corynebacte	647	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
575	33	58.9	440	2	Q5XK95_CANAL	Q5xk95 candida alb	648	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
576	33	58.9	440	2	Q59XE2_CANAL	Q59xe2 candida alb	649	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
577	33	58.9	442	2	Q4HBM9_9DEIO	Q4hbm9 deinococcus	650	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
578	33	58.9	443	2	Q7QKQ8_PSEF5	Q7qkq8 pseudomonas	651	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
579	33	58.9	446	2	Q4K6Q8_9DEIO	Q4k6q8 pseudomonas	652	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
580	33	58.9	447	2	Q4P9P8_USTMA	Q4p9p8 ustilago ma	653	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
581	33	58.9	447	2	Q4P9P8_USTMA	Q4p9p8 ustilago ma	654	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
582	33	58.9	448	2	Q9LJAE_9DELT	Q9ljae arabidopsis	655	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
583	33	58.9	450	2	Q6BIB0_DEBHA	Q6bib0 debaryomyce	656	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
584	33	58.9	451	2	Q6C655_YARLI	Q6c655 yarrowia li	657	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
585	33	58.9	451	2	Q522V1_MAGGR	Q522v1 magnaporthe	658	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
586	33	58.9	452	2	Q5FQI0_GLUOX	Q5fqio gluconobact	659	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
587	33	58.9	455	2	Q5DL07_9INFA	Q5dl07 influenza a	660	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
588	33	58.9	455	2	Q5DL08_9INFA	Q5dl08 influenza a	661	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
589	33	58.9	455	2	Q5DL09_9INFA	Q5dl09 influenza a	662	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
590	33	58.9	455	2	Q5DL09_9INFA	Q5dl09 influenza a	663	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
591	33	58.9	461	2	Q5DL10_9INFA	Q5dl10 influenza a	664	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
592	33	58.9	462	2	Q4PLH8_FUSPR	Q4plh8 fusarium pr	665	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
593	33	58.9	463	2	Q4I403_GIBZE	Q4i403 gibberella	666	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
594	33	58.9	466	2	Q6Y3Z8_TRYCR	Q6y3z8 trypanosoma	667	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
595	33	58.9	469	2	Q5NA57_ORYSA	Q5na57 oryza sativ	668	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
596	33	58.9	469	2	Q6GVY2_9INFA	Q6gv2 influenza a	669	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
597	33	58.9	469	2	Q6XV53_9INFA	Q6xv53 influenza a	670	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
598	33	58.9	469	2	Q6XV61_9INFA	Q6xv61 influenza a	671	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
599	33	58.9	472	2	Q6LW64_PHOPR	Q6lw64 photobacter	672	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
600	33	58.9	474	2	Q4RSP5_TETNG	Q4rsp5 tetradon n	673	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
601	33	58.9	474	2	Q87VQ8_PSESM	Q87vq8 pseudomonas	674	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
602	33	58.9	476	2	Q70FA5_9DELT	Q70fa5 mellittangiu	675	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
603	33	58.9	478	2	Q841L7_9ACTO	Q841l7 streptomyce	676	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
604	33	58.9	479	2	Q75BT3_ASHGO	Q75bt3 ashbya goss	677	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
605	33	58.9	481	2	Q4NRJ2_9DELT	Q4nrj2 anaeromyxob	678	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
606	33	58.9	486	2	Q8RSL8_9BACT	Q8rsl8 uncultured	679	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
607	33	58.9	519	2	Q6CS52_KLULA	Q6cs52 kluyveromyc	680	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
608	33	58.9	530	2	Q8NQV7_CORGL	Q8nqv7 corynebacte	681	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
609	33	58.9	532	2	Q5YRX1_NOCFA	Q5yrx1 nocardia fa	682	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
610	33	58.9	532	2	Q6BJ55_DEBHA	Q6bj55 debaryomyce	683	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
611	33	58.9	538	2	Q4N9C0_THEPA	Q4n9c0 theileria p	684	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
612	33	58.9	538	2	Q50491_STRCO	Q50491 streptomyce	685	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
613	33	58.9	541	2	MUTL_DEIRA	Q9rtt0 deinococcus	686	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
614	33	58.9	547	1	ML014_ARATH	Q94kbl1 arabidopsis	687	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce
615	33	58.9	554	1	ML014_ARATH	Q94kbl1 arabidopsis	688	33	58.9	610	2	Q6BMT2_DEBARYO	Q6bmt2 debaryomyce

689	33	58.9	858	2	Q8FCV2_ECOL6	Q8fcv2	escherichia	762	32	57.1	61	2	Q8GPI7_STAAL	Q8gpi7	staphylococ
690	33	58.9	858	2	Q6M0N7_BDBBA	Q6mqn7	bdellovibri	763	32	57.1	61	2	Q8NUB0_STAAN	Q8nue0	staphylococ
691	33	58.9	858	2	Q8X809_ECO57	Q8x809	escherichia	764	32	57.1	61	2	Q9AC61_STAAN	Q9ac61	staphylococ
692	33	58.9	866	2	Q55AA8_DICDI	Q55aa8	dictyosteli	765	32	57.1	61	2	Q6G5V6_STAAS	Q6g5v6	staphylococ
693	33	58.9	868	2	Q7X7Q3_ORISA	Q7x7q3	oryza sativ	766	32	57.1	62	2	Q6Z4T8_ORISA	Q6z4t8	oryza sativ
694	33	58.9	868	2	Q8A2H7_BACTN	Q8a2h7	bacteroides	767	32	57.1	67	2	Q5M2J0_SYNPE	Q5m2j0	synechococc
695	33	58.9	873	2	Q4W149_ASPFU	Q4w149	aspergillus	768	32	57.1	73	2	Q9KRI1_WIBCH	Q9kri1	vibrio chol
696	33	58.9	875	2	Q7S029_NEUCR	Q7s029	neurospora	769	32	57.1	73	2	Q92K84_RHIME	Q92k84	rhizobium m
697	33	58.9	881	2	Q4ZP54_PSESY	Q4zp54	pseudomonas	770	32	57.1	73	2	Q92K84_RHIME	Q92k84	rhizobium m
698	33	58.9	886	2	Q87X43_PRESM	Q87x43	pseudomonas	771	32	57.1	74	2	Q67MP1_SYMTH	Q67mp1	symbiobacte
699	33	58.9	896	2	Q7UYC9_RHOBA	Q7uyc9	rhodospirill	772	32	57.1	75	1	10KD_VIGUN	P8646	vigna ungui
700	33	58.9	901	1	MALT_ECOL6	Q87uyc9	rhodospirill	772	32	57.1	75	1	Q8CGL7_9HIST	Q8cgl7	trichys fas
701	33	58.9	901	1	MALT_ECOL6	Q8x701	escherichia	773	32	57.1	78	2	Q8EJD1_SHEON	Q8ejd1	shewanella
702	33	58.9	901	1	MALT_ECOL6	Q8cx55	escherichia	774	32	57.1	85	2	P72271_RHOER	P72271	rhodococcus
703	33	58.9	902	1	MALT_ECOL6	Q6G993	escherichia	775	32	57.1	87	2	Q81CW7_BACCR	Q81cw7	bacillus ce
704	33	58.9	906	2	Q9TTJ4_RABIT	P61hm8	photobacter	776	32	57.1	96	2	Q947F7_SOYBN	Q947f7	glycine max
705	33	58.9	908	2	Q7ZVL4_BRARE	Q9ttj4	oryctolagus	777	32	57.1	96	2	Q7TM50_ZZZZZ	Q7tm50	uncultured
706	33	58.9	933	2	Q4SZM3_TETNG	Q7zvl4	brachydanio	778	32	57.1	103	2	Q7TM02_ZZZZZ	Q7tm02	uncultured
707	33	58.9	934	2	Q4SN87_TETNG	Q4szm3	tetradon n	779	32	57.1	105	2	Q7VK37_HELHP	Q7vk37	helicobacte
708	33	58.9	939	2	Q4ZP59_PSESY	Q4sn87	tetradon n	780	32	57.1	106	2	Q4ZYT6_PSESY	Q4zyt6	pseudomonas
709	33	58.9	946	2	Q4NZU4_9DELTA	Q4zp59	pseudomonas	781	32	57.1	106	2	Q4IYJ6_AZOVI	Q4iyj6	azotobacter
710	33	58.9	952	2	Q8NIT0_NEUCR	Q4nzua	anaeromyxob	782	32	57.1	106	2	Q5PBT1_ANAMM	Q5pbt1	anaplasma m
711	33	58.9	973	2	Q6ZR04_HUMAN	Q8nit8	neurospora	783	32	57.1	110	2	Q73ZM0_MYCPA	Q73zm0	mycobacteri
712	33	58.9	982	2	Q7YZI1_BRARE	Q6zr04	homo sapien	784	32	57.1	111	2	Q9ALA7_SYNPF	Q9ala7	synechococc
713	33	58.9	987	2	Q7YZI1_BRARE	Q7zyz1	brachydanio	785	32	57.1	111	2	Q5N062_SYNPE	Q5n062	synechococc
714	33	58.9	1007	2	Q8C8T6_MOUSE	Q7zyz1	tetradon n	786	32	57.1	119	2	Q4KYM5_ORISA	Q4kym5	oryza sativ
715	33	58.9	1015	2	Q86JTP2_DICDI	Q8c8t6	mus muscucu	787	32	57.1	124	2	Q8MMV6_PIG	Q8mmv6	sus scrofa
716	33	58.9	1087	2	Q4HWP8_GIBZE	Q86jtp2	dictyosteli	788	32	57.1	124	2	Q852Y8_9CAUD	Q852y8	mycobacteri
717	33	58.9	1118	1	UBP8_HUMAN	Q4hwp8	gibberella	789	32	57.1	125	2	Q527J9_MAGGR	Q527j9	magnaporthe
718	33	58.9	1124	2	Q4Q685_LEIMA	P40818	homo sapien	790	32	57.1	131	2	Q58A12_CUCME	Q58a12	cucumis mel
719	33	58.9	1182	2	Q9Y7C1_MAGGR	Q4q685	leishmania	791	32	57.1	134	2	Q4KD10_PSEFS	Q4kd10	pseudomonas
720	33	58.9	1182	2	Q51TA9_MAGGR	Q9y7c1	magnaporthe	792	32	57.1	138	2	Q7SC72_NEUCR	Q7sc72	neurospora
721	33	58.9	1211	2	Q76YV0_9CAUD	Q51ta9	magnaporthe	793	32	57.1	143	2	Q9H8C4_HUMAN	Q9h8c4	homo sapien
722	33	58.9	1365	2	Q5AZI4_EMENI	Q76yv0	bacterioph	794	32	57.1	150	2	Q69K49_ORISA	Q69k49	oryza sativ
723	33	58.9	1426	2	Q522R0_MAGGR	Q5az14	aspergillus	795	32	57.1	156	2	Q6QRX5_9BETA	Q6qrx5	pongine her
724	33	58.9	1468	2	Q7XK59_ORISA	Q522r0	magnaporthe	796	32	57.1	158	2	Q912I0_PSEAE	Q912i0	pseudomonas
725	33	58.9	1468	2	Q8W5K9_ORISA	Q7xk59	oryza sativ	797	32	57.1	160	2	Q6MQ88_BDEBA	Q6mq88	bdellovibri
726	33	58.9	1483	2	Q5S864_CRYNE	Q8w5k9	oryza sativ	798	32	57.1	170	2	Q5P7T4_AZOSE	Q5p7t4	azocarcus sp
727	33	58.9	1483	2	Q5S864_CRYNE	Q5s864	cryptococcu	799	32	57.1	171	2	Q4L8M7_STAHT	Q4l8m7	staphylococ
728	33	58.9	1755	2	Q8CHC7_MOUSE	Q5rg17	cryptococcu	800	32	57.1	171	2	Q7UYX1_RHOBA	Q7uyx1	rhodospirill
729	33	58.9	2000	2	Q61N09_CABBR	Q8chc7	mus muscucu	801	32	57.1	172	2	Q88E67_PSEPK	Q88e67	pseudomonas
730	33	58.9	2068	2	Q572G1_9TRYP	Q61n09	caenorhabdi	802	32	57.1	174	2	Q8KJD2_RHILO	Q8kjd2	rhizobium l
731	33	58.9	2269	2	Q76K25_STRHA	Q572g1	trypanosoma	803	32	57.1	176	2	Q9ZGS2_ECO57	Q9zgs2	escherichia
732	33	58.9	2333	2	Q5SNZ5_MOUSE	Q76k25	streptomyce	804	32	57.1	177	2	Q5M9B6_XENLA	Q5m9b6	xenopus lae
733	33	58.9	2333	2	Q6RKE8_COCHS	Q5swz5	mus muscucu	805	32	57.1	178	2	Q6Z462_ORISA	Q6z462	oryza sativ
734	33	58.9	3362	2	Q76K24_STRHA	Q6rke8	cochlobolu	806	32	57.1	179	2	Q4KAB0_PSEFS	Q4kab0	pseudomonas
735	33	58.9	3528	2	Q5Y9G4_9ACTO	Q76k24	streptomyce	807	32	57.1	182	2	Q6I133_DROME	Q6i133	grosophila
736	33	58.9	5826	2	Q76KY0_STRHA	Q5y9g4	aeromicrobi	808	32	57.1	182	2	Q62373_CAEBL	Q62373	caenorhabdi
737	33	58.9	6521	2	Q8RL72_PSEFL	Q76ky0	streptomyce	809	32	57.1	182	2	Q4MIW2_BACCE	Q4miw2	bacillus ce
738	32.5	58.0	79	2	Q9AC49_CAUCR	Q8rl72	pseudomonas	810	32	57.1	182	2	Q4MIW2_BACCE	Q4miw2	bacillus ce
739	32.5	58.0	88	2	Q914K3_9VIRU	Q9ac49	caulobacter	811	32	57.1	182	2	Q6HI77_BACHK	Q6hi77	bacillus th
740	32.5	58.0	130	2	Q8FAB0_ECOL6	Q914k3	sulfolobus	812	32	57.1	182	2	Q737I1_BACCI	Q737i1	bacillus ce
741	32.5	58.0	176	2	Q7TVD9_MYCBO	Q8fab0	escherichia	813	32	57.1	182	2	Q81P29_BACAN	Q81p29	bacillus an
742	32.5	58.0	176	2	Q05443_MYCTU	Q7tvd9	mycobacteri	814	32	57.1	186	2	Q6LG71_PROPR	Q6lg71	photobacter
743	32.5	58.0	188	2	Q4TB10_TETNG	O05443	mycobacteri	815	32	57.1	186	2	Q8Y894_ANASP	Q8y894	anabaena sp
744	32.5	58.0	246	2	Q6LGC3_PROPR	Q4tb10	tetradon n	816	32	57.1	191	2	Q64UN9_BACPR	Q64un9	bacteroides
745	32.5	58.0	343	2	Q6LRFP_PROPR	Q6lgc3	photobacter	817	32	57.1	194	2	Q9RX50_DEIRA	Q9rx50	deinococcus
746	32.5	58.0	345	2	Q6KD97_ECOL6	Q6lrfp	photobacter	818	32	57.1	195	2	Q5I081_XENTR	Q5i081	xenopus tro
747	32.5	58.0	352	2	Q6FIW8_ECOL6	Q6kd97	escherichia	819	32	57.1	197	2	Q58353_PYRHO	Q58353	pyrococcus
748	32.5	58.0	355	2	Q6LHY5_PROPR	Q6fiw8	escherichia	820	32	57.1	198	1	COMD9_HUMAN	Q8p000	homo sapien
749	32.5	58.0	398	2	Q84701_CHLTR	Q6lhy5	photobacter	821	32	57.1	198	1	COMD9_MOUSE	Q8k2q0	mus muscucu
750	32.5	58.0	402	2	Q7VXD5_BORPE	Q84701	chlamydia t	822	32	57.1	198	1	COMD9_MOUSE	Q8k2q0	mus muscucu
751	32.5	58.0	402	2	Q7W6N8_BORBA	Q7vxd6	bordetella	823	32	57.1	200	2	Q5TYH1_ANOGA	Q5tyh1	anopheles g
752	32.5	58.0	402	2	Q7WHM1_BORBA	Q7w6n8	bordetella	824	32	57.1	203	2	Q8GWR6_9ATHX	Q8gwr6	arabidopsis
753	32.5	58.0	567	1	FIG1_HUMAN	Q7whm1	bordetella	825	32	57.1	204	2	Q6F443_PLUXY	Q6f443	plutella xy
754	32.5	58.0	575	2	Q4GZNI_HUMAN	Q9grq9	homo sapien	826	32	57.1	205	2	Q4NNN3_9DELT	Q4nnn3	anaeromyxob
755	32.5	58.0	575	2	Q4GZNI_HUMAN	Q4gzn1	homo sapien	827	32	57.1	205	2	Q82PS6_STRAW	Q82ps6	streptomyce
756	32.5	58.0	588	2	Q8TEM3_HUMAN	Q4gzn1	homo sapien	828	32	57.1	206	2	Q6ZHF1_ORISA	Q6zhf1	oryza sativ
757	32.5	58.0	588	2	Q6P2Q3_HUMAN	Q8tem5	homo sapien	829	32	57.1	206	2	Q7UR11_RHOBA	Q7ur11	rhodospirall
758	32.5	58.0	919	2	Q6CW36_KLUHA	Q6p2q3	homo sapien	830	32	57.1	210	2	Q53NC9_ORISA	Q53nc9	oryza sativ
759	32.5	58.0	1160	2	Q88D24_PSEPK	Q6cw36	kluveromyce	831	32	57.1	210	2	Q5N213_SYNPE	Q5n213	synechococc
760	32.5	58.0	1186	2	Q5Z3U8_NOCFA	Q88d24	pseudomonas	832	32	57.1	213	2	Q677F9_9ASPA	Q677f9	hyacinthos
761	32.5	58.0	1904	2	Q41L82_GIBZE	Q5z3u8	nocardia fa	833	32	57.1	216	2	Q7Z6R3_HUMAN	Q7z6r3	homo sapien
						Q41l82	gibberella	834	32	57.1	221	2	Q6A970_PROAC	Q6a970	propionibac

835	32	57.1	223	2	Q412B5_GIBZ2	Q412e5 gibberella	908	32	57.1	324	2	Q9YEP2_AERPE	Q9yep2 aeropyrum p
836	32	57.1	223	2	Q5Z5X5_ORYSA	Q5z5x5 oryza sativ	909	32	57.1	326	2	Q7USG3_RHOBA	Q7usg3 rhodopirell
837	32	57.1	223	2	Q3YV80_NOCFA	Q3yvs0 nocardia fa	910	32	57.1	326	2	Q5SIG8_THET8	Q5sig8 thermus the
838	32	57.1	224	2	Q5ZF83_PLAMU	Q5zf83 plantago ma	911	32	57.1	327	2	Q72IU7_THET2	Q72iu7 thermus the
839	32	57.1	224	2	Q711G5_TOBAC	Q711g5 nicotiana t	912	32	57.1	327	2	Q7VZS8_BORPE	Q7vzs8 bordetella
840	32	57.1	226	2	Q86SR2_HUMAN	Q86sr2 homo sapien	913	32	57.1	327	2	Q7W583_BORPA	Q7w583 bordetella
841	32	57.1	228	2	Q9CNM1_PASMU	Q9cnm1 pasteurella	914	32	57.1	327	2	Q7WCR3_BORDE	Q7wcr3 bordetella
842	32	57.1	228	2	Q82LX4_STRAW	Q82lx4 streptomyce	915	32	57.1	328	2	Q60DR3_ORYSA	Q60dr3 oryza sativ
843	32	57.1	229	2	Q4J663_AZOVI	Q4j663 azotobacter	916	32	57.1	328	2	Q6A8Q0_PROAC	Q6a8q0 propionibac
844	32	57.1	230	2	Q3SSW8_HUMAN	Q3ssw8 homo sapien	917	32	57.1	330	2	Q7WP27_BORBR	Q7wp27 bordetella
845	32	57.1	230	2	Q3SIV6_ARATH	Q3siv6 arabidopsis	918	32	57.1	330	2	SPRML1_RAT	P56225 rattus norv
846	32	57.1	230	2	Q3ZB92_RHOER	Q3zb92 rhodococcus	919	32	57.1	335	1	Q75QH4_CAPCH	Q75qh4 capsicum ch
847	32	57.1	230	2	Q7NS28_CHRVO	Q7ns28 chromobacte	920	32	57.1	335	2	Q8ELP2_OCEIH	Q8elp2 oceanobacil
848	32	57.1	237	2	Q5SSX0_HUMAN	Q5ssx0 homo sapien	921	32	57.1	337	2	Q4HU90_GIBZE	Q4hu90 gibberella
849	32	57.1	239	2	Q6DH54_BRARE	Q6dh54 brachydanio	922	32	57.1	340	2	Q4HU90_GIBZE	Q4hu90 gibberella
850	32	57.1	243	2	Q5TY12_ANOGA	Q5ty12 anopheles g	923	32	57.1	340	2	Q8P152_XANAC	Q8p152 xanthomonas
851	32	57.1	245	2	Q4NI84_9MICC	Q4ni84 arthrobacte	924	32	57.1	340	2	Q95VR4_TRIVA	Q95vr4 trichomonas
852	32	57.1	246	2	Q8Y0L4_RALSO	Q8y0l4 ralstonia s	925	32	57.1	341	2	Q95VR4_TRIVA	Q95vr4 trichomonas
853	32	57.1	247	2	Q4TP77_9PHN	Q4tp77 erythrobact	926	32	57.1	341	2	Q4HAN3_9DEIO	Q4han3 deinococcus
854	32	57.1	250	2	Q7NG30_GLOVI	Q7ng30 gloeobacter	927	32	57.1	344	2	Q4USX3_XANCP	Q4usx3 xanthomonas
855	32	57.1	250	2	Q82PG7_STRAW	Q82pg7 streptomyce	928	32	57.1	344	2	Q6AL73_DESPS	Q6al73 desulfotale
856	32	57.1	251	1	CAP7_HUMAN	P20160 homo sapien	929	32	57.1	344	2	Q7NEH3_GLOVI	Q7neh3 gloeobacter
857	32	57.1	251	2	Q52LG4_HUMAN	Q52lg4 homo sapien	930	32	57.1	345	2	Q8PAP2_XANCP	Q8pap2 xanthomonas
858	32	57.1	253	2	Q6YI13_ORYSA	Q6yvi3 oryza sativ	931	32	57.1	345	2	Q61IX5_CAEBR	Q61ix5 caenorhabdi
859	32	57.1	253	2	Q89GZ4_BRAJA	Q89gz4 bradyrhizob	932	32	57.1	345	2	Q9CNK9_PASMU	Q9cnk9 pasteurella
860	32	57.1	254	2	Q88YF6_LACPL	Q88yf6 lactobacill	933	32	57.1	346	2	Q24548_VIGUN	Q24548 vigna ungui
861	32	57.1	255	2	Q7SE51_NEUCR	Q7se51 neurospora	934	32	57.1	347	2	Q67PH3_SYMTH	Q67ph3 symbiobacte
862	32	57.1	255	2	Q6N9Q3_RHOPA	Q6n9q3 rhodopseudo	935	32	57.1	347	2	Q82705_ORYSA	Q82705 oryza sativ
863	32	57.1	256	2	Q9KX37_9RHIZ	Q9kx37 methylocyst	936	32	57.1	349	2	Q9HRH2_AZOVI	Q9hrh2 azotobacter
864	32	57.1	256	2	Q7NUF0_CHRVO	Q7nuf0 chromobacte	937	32	57.1	349	2	Q4IY49_AZOVI	Q4iy49 azotobacter
865	32	57.1	260	2	Q568S5_BRARE	Q568s5 brachydanio	938	32	57.1	350	2	Q5L4X0_CHLAB	Q5l4x0 chlamydophi
866	32	57.1	261	2	Q92TL1_RHIME	Q92tl1 rhizobium m	939	32	57.1	351	2	Q6C6P0_YARLI	Q6c6p0 yarrowia li
867	32	57.1	265	1	FAB1_PSEAE	Q9zfe4 pseudomonas	940	32	57.1	351	2	Q4RG21_TETNG	Q4rg21 tetraodon n
868	32	57.1	267	2	Q5TXH7_ANOGA	Q5txh7 anopheles g	941	32	57.1	351	2	Q82KS5_SALTY	Q82ks5 salmonella
869	32	57.1	268	2	Q52BL3_MAGGR	Q52bl3 magnaporthe	942	32	57.1	352	2	Q69F96_PHAVU	Q69f96 pseudomonas
870	32	57.1	268	2	P93332_MEDTR	P93332 medicago tr	943	32	57.1	353	2	Q4KKQ8_PSEF5	Q4kkq8 pseudomonas
871	32	57.1	270	2	Q4M0V0_9BURK	Q4m0v0 burkholderi	944	32	57.1	353	2	Q8MHQ3_CHICK	Q8mhq3 gallus gall
872	32	57.1	272	2	Q6X3Z0_CHICK	Q6x3z0 gallus gall	945	32	57.1	353	2	Q9BCW3_CHICK	Q9bcw3 gallus gall
873	32	57.1	272	2	Q6X3Z1_CHICK	Q6x3z1 gallus gall	946	32	57.1	355	2	Q5SKQ8_THET8	Q5skq8 thermus the
874	32	57.1	273	2	Q4V0D2_XANCP	Q4v0d2 xanthomonas	947	32	57.1	355	2	Q72L47_THET2	Q72l47 thermus the
875	32	57.1	273	2	Q8PE38_XANCP	Q8pe38 xanthomonas	948	32	57.1	355	2	O22427_ARATH	O22427 arabidopsis
876	32	57.1	278	2	Q6C344_YARLI	Q6c344 yarrowia li	949	32	57.1	357	2	Q50HT9_ARATH	Q50ht9 arabidopsis
877	32	57.1	282	2	Q89HL6_BRAJA	Q89hl6 bradyrhizob	951	32	57.1	357	2	Q50HV8_WHEAT	Q50hv8 triticum ae
878	32	57.1	282	2	Q89PW0_BRAJA	Q89pw0 bradyrhizob	952	32	57.1	357	2	Q7XB44_HORVU	Q7xb44 hordeum vul
879	32	57.1	283	2	Q5W495_STRT2	Q5w495 streptococc	953	32	57.1	357	2	Q4KE51_PSEF5	Q4ke51 pseudomonas
880	32	57.1	284	2	Q56WY4_ARATH	Q56wy4 arabidopsis	954	32	57.1	359	2	Q8VWN8_GOSHI	Q8vwn8 gosypium h
881	32	57.1	284	2	Q6N287_RHOPA	Q6n287 rhodopseudo	955	32	57.1	360	2	Q8LB19_ARATH	Q8lb19 arabidopsis
882	32	57.1	285	2	Q33629_SACER	Q33629 saccharopol	956	32	57.1	360	2	Q9LFW1_ARATH	Q9lfw1 arabidopsis
883	32	57.1	285	2	Q4KHK1_PSEF5	Q4khh1 pseudomonas	957	32	57.1	360	2	Q4LZF0_9BURK	Q4lzf0 burkholderi
884	32	57.1	286	2	Q51J97_MAGGR	Q51j97 magnaporthe	958	32	57.1	360	2	Q82H28_STRAW	Q82h28 streptomyce
885	32	57.1	286	2	Q5A6L2_CAUCR	Q5a6l2 caulobacter	959	32	57.1	360	2	Q891D2_BRAJA	Q891d2 bradyrhizob
886	32	57.1	287	2	Q58DH6_BOVIN	Q58dh6 bos taurus	960	32	57.1	361	2	Q762H2_RHORH	Q762h2 rhodococcus
887	32	57.1	288	2	Q8GH80_9ACTO	Q8gh80 magnaporthe	961	32	57.1	361	2	Q762H2_RHORH	Q762h2 rhodococcus
888	32	57.1	288	2	Q51SY8_MAGGR	Q51sy8 streptomyce	962	32	57.1	361	2	Q4NJZ7_9MICC	Q4njz7 artibacte
889	32	57.1	289	2	Q4LT59_9BURK	Q4lt59 burkholderi	963	32	57.1	362	2	O22666_ARATH	O22666 arabidopsis
890	32	57.1	289	2	Q82AZ7_STRAW	Q82az7 streptomyce	964	32	57.1	362	2	Q61V07_LYCES	Q61v07 lycopersico
891	32	57.1	290	2	Q5CMC7_CRYHO	Q5cmc7 cryptospori	965	32	57.1	362	2	Q9XE81_SORBI	Q9xe81 sorghum bic
892	32	57.1	291	2	Q58AG5_9BURK	Q58ag5 ralstonia m	966	32	57.1	363	2	Q6MBF3_PARUW	Q6mbf3 parschlamyd
893	32	57.1	293	2	Q7Z5A4_HUMAN	Q7z5a4 homo sapien	967	32	57.1	364	1	UPTG_MAIZE	P80607 zea mays lm
894	32	57.1	296	2	Q4M0B1_9BURK	Q4m0b1 burkholderi	968	32	57.1	364	1	UPTG_PEA	O04300 pisum sativ
895	32	57.1	296	2	Q7NUS2_CHRVO	Q7nus2 chromobacte	969	32	57.1	364	2	O22428_ARATH	O22428 arabidopsis
896	32	57.1	299	2	Q61CS5_CAEBR	Q61cs5 caenorhabdi	970	32	57.1	364	2	Q8H8T0_ORYSA	Q8ht0 oryza sativ
897	32	57.1	299	2	Q83B48_COXBU	Q83b48 coxiella bu	971	32	57.1	364	2	Q9FUN9_ORYSA	Q9fun9 oryza sativ
898	32	57.1	301	2	Q9HS99_HALSA	Q9hs99 halobacteri	972	32	57.1	364	2	Q9Z835_ORYSA	Q9z835 oryza sativ
899	32	57.1	301	2	Q51728_ENTHI	Q51728 entamoeba h	973	32	57.1	364	2	Q8B549_VIBVU	Q8b549 vibrio vuln
900	32	57.1	302	2	Q5H2Y7_XANOR	Q5h2y7 xanthomonas	974	32	57.1	365	1	Q7MBY6_VIBVY	Q7mby6 vibrio vuln
901	32	57.1	308	2	Q7NKJ2_CHRVO	Q7nkj2 chromobacte	975	32	57.1	365	2	UPTG1_SOLTU	UPTG1 solanum tub
902	32	57.1	311	2	Q8FWZ7_BRUSU	Q8fwz7 bruceella su	976	32	57.1	366	1	UPTG2_SOLTU	UPTG2 solanum tub
903	32	57.1	312	2	Q5TRY8_ANOGA	Q5try8 anopheles g	977	32	57.1	366	2	Q6Z4G3_ORYSA	Q6z4g3 oryza sativ
904	32	57.1	317	2	Q9RJ64_STRCO	Q9rj64 streptomyce	978	32	57.1	366	2	Q57HH7_SALCH	Q57hh7 salmonella c
905	32	57.1	318	2	Q61826_CAEBL	Q61826 caenorhabdi	979	32	57.1	366	2	O84715_CHLTR	O84715 chlamydia c
906	32	57.1	320	2	Q5V0P1_HALMA	Q5v0p1 haloarcula	980	32	57.1	366	2	Q5PKF8_SALPA	Q5pkf8 salmonella
907	32	57.1	322	2	O24547_VIGUN	O24547 vigna ungui							

```
981 32 57.1 366 2 Q821M3 chlamydom
982 32 57.1 366 2 Q9PL17 chlamydia m
983 32 57.1 366 2 Q822U9 salmonella
984 32 57.1 366 2 Q822U9 salm
985 32 57.1 366 2 Q822U9 salm
986 32 57.1 366 2 Q822U9 salm
987 32 57.1 366 2 Q822U9 salm
988 32 57.1 366 2 Q822U9 salm
989 32 57.1 366 2 Q822U9 salm
990 32 57.1 366 2 Q822U9 salm
991 32 57.1 366 2 Q822U9 salm
992 32 57.1 366 2 Q822U9 salm
993 32 57.1 366 2 Q822U9 salm
994 32 57.1 366 2 Q822U9 salm
995 32 57.1 366 2 Q822U9 salm
996 32 57.1 366 2 Q822U9 salm
997 32 57.1 366 2 Q822U9 salm
998 32 57.1 366 2 Q822U9 salm
999 32 57.1 366 2 Q822U9 salm
1000 32 57.1 366 2 Q822U9 salm
```

ALIGNMENTS

```
RESULT 1
Q4JY45 CORJK PRELIMINARY; PRT; 242 AA.
ID Q4JY45 CORJK PRELIMINARY; PRT; 242 AA.
AC Q4JY45;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE MrA protein (EC 1.8.4.6).
GN Name=MrA; ORFNames=jk0110;
OS Corynebacterium jeikeium (strain K411).
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=306537;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=K411;
RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
RA Albersmeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
RA Kalinowski J., Meyer F., Rupp O., Schmelzer S., Vienne P.,
RA Puhler A.;
RT "Complete Genome Sequence and Analysis of the Multiresistant
RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requirement
RT Bacterium of the Human Skin Flora.";
RL J. Bacteriol. 187:4671-4682(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=K411;
RA Linke B., Tauch A.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RW EMBL; CR91997; CA16262.1; -; Genomic_DNA.
KW Oxidoreductase.
SQ SEQUENCE 242 AA; 26352 MW; B1776A45CDAE8D3E CRC64;
```

```
Query Match 76.8%; Score 43; DB 2; Length 242;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GRVRDQVAGW 10
    |||||
Db 171 GRVRQAIAGW 180
```

```
RESULT 2
Q8RSP4 RHILV PRELIMINARY; PRT; 467 AA.
ID Q8RSP4 RHILV PRELIMINARY; PRT; 467 AA.
AC Q8RSP4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FixW3 protein.
GN Name=fixW3;
OS Rhizobium leguminosarum (biovar viciae).
OG Plasmid pRL6J1.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=387;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15375140; DOI=10.1128/JB.186.19.6586-6594.2004;
RA Martinez M., Palacios J.M., Imperial J., Ruiz-Argueso T.;
RT "Symbiotic autoregulation of nifA expression in Rhizobium
RT leguminosarum bv. viciae";
RL J. Bacteriol. 186:6586-6594(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Martinez M.;
RT "Regulation simbiótica de la expresión del sistema hidrogenasa por
RT NifA en R. leguminosarum bv. viciae.";
RL Thesis (2000), Department of Biología, Universidad Politécnica
RL Madrid, Madrid, Spain.
DR EMBL; AJ431176; CAD24022.1; -; Genomic_DNA.
KW Plasmid.
SQ SEQUENCE 467 AA; 52042 MW; 412603D21D20AAA7 CRC64;
```

```
Query Match 73.2%; Score 41; DB 2; Length 467;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 GRVRDQVAGW 10
    |||||
Db 323 GRMRDMDAGW 332
```

```
RESULT 3
Q4TM05 9SPHN PRELIMINARY; PRT; 776 AA.
ID Q4TM05 9SPHN PRELIMINARY; PRT; 776 AA.
AC Q4TM05;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative cation transporting P-type ATPase.
GN ORFNames=ELI2122;
OS Erythrobacter litoralis HTCC2594.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Erythrobacteraceae; Erythrobacter.
OX NCBI_TaxID=314225;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HTCC2594;
RA Giovannoni S.J., Cho J.-C., Ferreira S., Johnson J., Kravitz S.,
RA Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
RA Friedman R., Venter J.C.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
CC -!- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; NAG01000006; EAL74862.1; -; Genomic_DNA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB_Hv.
DR InterPro; IPR001757; ATPase-E1-E2.
DR InterPro; IPR001756; Cu ATPase.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR007029; YHS.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR Pfam; PF04945; YHS; 1.
```

```

DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00943; CUATPASE.
DR TIGRFAMs; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB1_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase P-type; 2.
DR PROSITE; PS00154; ATPase E1_E2; UNKNOWN_1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 776 AA; 81530 MW; 228466303B32A46 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 776;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: |||||
Db 371 RLADQVAGW 379

RESULT 4
P74512 SYN33
ID P74512 SYN33 PRELIMINARY; PRT; 780 AA.
AC P74512;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cation-transporting ATPase; E1-E2 ATPase.
GN OrderedLocustNames=slr1950;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; BA000022; BAA18616.1; -; Genomic_DNA.
DR FIR; S76487; S76487.
DR HSSP; O32220; 10Q3.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004008; F:copper-exporting ATPase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0006825; P:copper ion transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB1_hvy.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR001877; Cu ATPase.
DR InterPro; IPR005834; Dehal-like hydro.
DR InterPro; IPR008250; E1-E2 ATPase reg.
DR InterPro; IPR006121; HeavyMe transpt.
DR InterPro; IPR000150; Hypothes cof.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00942; CUATPASE.
DR TIGRFAMs; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB1_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase P-type; 2.
DR PROSITE; PS00154; ATPase E1_E2; UNKNOWN_1.

DR PRINTS; PR001229; COF_2; UNKNOWN_1.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS00846; HMA_2; 1.
KW Complete proteome.
SQ SEQUENCE 780 AA; 82571 MW; 5BC2B2573124708C CRC64;

Query Match 73.2%; Score 41; DB 2; Length 780;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|: |||||
Db 357 RLADQVAGW 365

RESULT 5
Q89WI7 BRAJA
ID Q89WI7 BRAJA PRELIMINARY; PRT; 823 AA.
AC Q89WI7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocustNames=b10700;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484988; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -! SURCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -! SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; BA000040; BAC45965.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004008; F:copper-exporting ATPase activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0006825; P:copper ion transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR006403; ATPase-IB1_Cu.
DR InterPro; IPR006416; ATPase-IB1_hvy.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR001756; Cu ATPase.
DR InterPro; IPR005834; Dehal-like hydro.
DR InterPro; IPR008250; E1-E2 ATPase reg.
DR InterPro; IPR012348; Ribncl_rdxase_rel.
DR InterPro; IPR011017; TRASH.
DR InterPro; IPR007029; YHS.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR Pfam; PF04945; YHS; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00943; CUATPASE.
DR SMART; SM00746; TRASH; 1.
DR TIGRFAMs; TIGR01511; ATPase-IB1_Cu; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB1_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase P-type; 1.
DR PROSITE; PS00154; ATPase E1_E2; UNKNOWN_1.

```

KW ATP-binding; Complete proteome; Hydrolase; Nucleotide-binding;
KW Transmembrane.
SQ SEQUENCE 823 AA; 86327 MW; 2D5E3A01D6CF06 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 823;
Best Local Similarity 77.8%; Pred. No. 99;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|:|||||
DB 420 RLADQVAGW 428

RESULT 6

ID Q9KZY7_STRCO PRELIMINARY; PRT; 832 AA.
AC Q9KZY7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative D-alanyl-D-alanine carboxypeptidase.
GN OrderedLocusNames=SC04439; ORFNames=SCD6.17c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939120; CAB88446.1; -; Genomic_DNA.
DR HSSP; P39042; 1J9M.
DR GO; GO:0004185; F:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001967; Peptidase_S11.
DR Pfam; PF00768; Peptidase_S11; 1.
DR PRINTS; PR00725; DADACBPTASE1.
KW Carboxypeptidase; Complete proteome.
SQ SEQUENCE 832 AA; 84223 MW; A02645EBDC63EF47 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 832;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|||||
DB 79 GRLRDAVAW 88

RESULT 7

ID Q7VLA1_HAEDU PRELIMINARY; PRT; 130 AA.
AC Q7VLA1;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=HDI1563;
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.

OX NCBI_TaxID=730;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017155; AAP96347.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 130 AA; 14535 MW; C45D929F53C66461 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 130;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|||||
DB 80 GKIREQLMGW 89

RESULT 8

Q9RL18_STRCO
ID Q9RL18_STRCO PRELIMINARY; PRT; 294 AA.
AC Q9RL18;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative oxidoreductase.
GN OrderedLocusNames=SC00299; ORFNames=SC5G9.08;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB55656.1; -; Genomic_DNA.
DR HSSP; P06632; 1HW6.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001395; Aldo/ket red.
DR Pfam; PF00248; Aldo_ket_red; 1.
DR PRINTS; PR00069; ALDKETEDTASE.
DR ProDom; PD000288; Aldo/ket_red; 1.
KW Complete proteome.
SQ SEQUENCE 294 AA; 30906 MW; 75DDE0746A1B3EF5 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 294;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|:|||||
DB 97 GAVRDQCGW 106

RESULT 9

Q7RL87_PLAYO
ID Q7RL87_PLAYO PRELIMINARY; PRT; 302 AA.
AC Q7RL87;

Query Match	71.4%;	Score 40;	DB 2;	Length 313;	
Best Local Similarity	77.8%;	Pred. No. 58;			
Matches	7;	Conservative	0;	Mismatches	2; Indels 0; Gaps 0;

QY	2	RVRDQVAGW	10		
DB	107	RPRDQVVGW	115		

RESULT 11	
Q4XWR8_PLACH	
ID	Q4XWR8_PLACH PRELIMINARY; PRT; 315 AA.
AC	Q4XWR8
DT	13-SEP-2005 (TREMBlrel. 31, Created)
DT	13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT	13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE	Hypothetical protein.
GN	ORFNames=PC001205.02.0;
OS	Plasmodium chabaudi.
OC	Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5825;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA	Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA	James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA	Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA	Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA	Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,
RT	"A comprehensive survey of the Plasmodium life cycle by genomic,
RT	transcriptomic, and proteomic analyses.";
RL	Science 307:82-86 (2005).
CC	-1- CAUTION: The sequence shown here is derived from an
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
DR	EMBL; CAAT01002585; CAH78643.1; -; Genomic_DNA.
KW	Hypothetical protein.
SQ	SEQUENCE 315 AA; 36144 MW; B12CA075E710D75A CRC64;

Query Match	71.4%;	Score 40;	DB 2;	Length 315;	
Best Local Similarity	77.8%;	Pred. No. 58;			
Matches	7;	Conservative	0;	Mismatches	2; Indels 0; Gaps 0;

QY	2	RVRDQVAGW	10		
DB	107	RPRDQVVGW	115		

RESULT 12	
Q812X0_PLAF7	
ID	Q812X0_PLAF7 PRELIMINARY; PRT; 320 AA.
AC	Q812X0;
DT	01-MAR-2003 (TREMBlrel. 23, Created)
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE	Hypothetical protein PF10895c.
GN	Name=PF10895c;
OS	Plasmodium falciparum (isolate 3D7).
OC	Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=36329;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RX	MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA	Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA	Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA	Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA	Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA	Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA	Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA	Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA	Humphray S., Jagels K., James K.D., Johnson D., Kethornou A.,
RA	Knight A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

```
RA Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajadream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sultston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002);
DR EMBL; AL929357; CAD51865.1; -; Genomic_DNA.
DR InterPro; IPR003639; MOV34-1.
DR InterPro; IPR000555; MOV34_MPN_PAD1.
DR Pfam; PF01398; MOV34_1.
DR ProDom; PD363422; MOV34-1; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Hypothetical protein.
SQ SEQUENCE 320 AA; 36820 MW; 158AA64F9BC5F8A4 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 320;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 RVRDQVAGW 10
| | | | |
Db 107 RPRDQVGVW 115

RESULT 13
Q7RXK6 NEUCR PRELIMINARY; PRT; 369 AA.
AC Q7RXK6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein.
GN Name=NCU00469.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kanal M., Kamysasellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Navig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nubbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000735; EAA27552.1; -; Genomic_DNA.
DR InterPro; IPR000408; Reg_chr_condens.
DR Pfam; PF00415; RCC1_1.
DR PRINTS; PR00633; RCCNDNSATTON.
DR PROSITE; PS00012; RCC1_3; 4.
KW Hypothetical protein.
SQ SEQUENCE 369 AA; 38697 MW; 0D83BACF5E420C20 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 369;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Cy 1 GRVRDQVAGW 10
| | | | |
Db 245 GEVPDQVGVW 254

RESULT 14
Q7TLZ4_9ZZZZ PRELIMINARY; PRT; 98 AA.
AC Q7TLZ4_9ZZZZ;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative xylanase (Fragment).
OS uncultured organism.
OC unclassified; environmental samples.
OX NCBI_TaxID=155900;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Radomski C.C.A., Seow K.T., Warren R.A.J., Yap W.H.;
RT "Method for isolating xylanase gene sequences from soil DNA,
RT compositions useful in such method and compositions obtained
RT thereby.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Chittaranjan S., Radomski C.C., Chow M.L., Davies J., Axelrood P.E.;
RT "Retrieval of novel DNA sequences encoding xylanase-like genes from
RT British Columbia forest soils.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427439; AAP87537.1; -; Genomic_DNA.
DR HSSP; P26514; 1BOW.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001000; Glyco_hydro_10.
DR Pfam; PF00331; Glyco_hydro_10; 1.
KW Glycosidase; Hydrolase; Xylan degradation.
FT NON_TER 1 1
FT NON_TER 98 98
SQ SEQUENCE 98 AA; 10854 MW; 4142C83DD2E89A70 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 98;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Cy 1 GRVRDQVAGW 10
| | | | |
Db 38 GRYRGRIAGW 47

RESULT 15
Q9FRG1_ORYSA PRELIMINARY; PRT; 124 AA.
AC Q9FRG1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Expressed protein.
GN Name=Oe03g47210;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNB0009F04 genomic sequence.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
```

```
RA Buell R.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079930; AAC46067.1; -, Genomic_DNA.
DR Gramene; Q9FRG1; -.
SQ SEQUENCE 124 AA; 12612 MW; 87AACB98AC6C2E78 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 124;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db 35 GRPLDRVAGW 44
|||:|||||
|||:|||||

RESULT 16
ID HIS1 CAUCR STANDARD; PRT; 320 AA.
AC Q9A2P5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ATP phosphoribosyltransferase (EC 2.4.2.17) (ATP-PRTase) (ATP-PRT).
GN Name=hisG; OrderedLocusNames=CC3511;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
STRAIN=ATCC 19089 / CB15;
MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Catalyzes the condensation of ATP and PRPP to form N'-
CC 5'-phosphoribosyl-ATP (PR-ATP). Has a crucial role in the pathway
CC because the rate of histidine biosynthesis seems to be controlled
CC primarily by regulation of hisG enzymatic activity (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 1-(5-phospho-D-riboseyl)-ATP + diphosphate =
CC ATP + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- ENZYME REGULATION: Feedback inhibited by histidine (By
CC similarity).
CC -1- PATHWAY: Amino-acid biosynthesis; L-histidine biosynthesis; L-
CC histidine from PRPP; step 1.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.
CC Long subfamily.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; AE006009; AAK25473.1; -, Genomic_DNA.
DR PIR; E87684; E87684.
DR TIGR; CC3511; -.
DR HAMAP; MF_00079; -, 1.
DR InterPro; IPR001348; ATP_phepho_trans.
DR Pfam; PF01634; HisG; 1.
DR ProDom; PD003516; ATP_phepho_trans; 1.
DR PROSITE; PS01316; ATP_P_PHORIBOSYLTR; 1.
KW Amino-acid biosynthesis; Complete proteome; Glycosyltransferase;
```

```
KW Histidine biosynthesis; Magnesium; Metal-binding; Transferase.
SQ SEQUENCE 320 AA; 33587 MW; B3126BC54400A6C1 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 320;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db 13 GRUKDQVEAW 22
|||:|||||
|||:|||||

RESULT 17
ID Q4LRX7 9BURK PRELIMINARY; PRT; 334 AA.
AC Q4LRX7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Regulatory protein, LysR:lysR, substrate-binding.
GN ORFNames=Bcen2424DRAFT_3611;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331272;
[1]
NUCLEOTIDE SEQUENCE.
STRAIN=HI2424;
US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
STRAIN=HI2424;
US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AALH0100022; EAM18796.1; -, Genomic DNA.
SQ SEQUENCE 334 AA; 35491 MW; 0D08091E13868B4E CRC64;

Query Match 69.6%; Score 39; DB 2; Length 334;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
Db 254 GRLEVLGAW 263
|||:|||||
|||:|||||

RESULT 18
Q8YV68 ANASP PRELIMINARY; PRT; 360 AA.
ID Q8YV68 ANASP PRELIMINARY;
AC Q8YV68;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE All0983 protein.
GN OrderedLocusNames=all0983;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
```

```
RA Matanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR ENBL; BA000019; BAB72940.1; -; Genomic_DNA.
DR PIR; AD1929; AD1929.
DR HSP; P05806; INPC.
DR GO; GO:000576; C:extracellular region; IEA.
DR GO; GO:0004222; P:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001570; peptidase_M4.
DR InterPro; IPR006025; pept_M_Zn_BS.
DR Pfam; PF01447; Peptidase_M4; 1.
DR Pfam; PF02868; Peptidase_M4_C; 1.
DR PRINTS; PRO0730; THERMOLYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 360 AA; 39795 MW; D14F84AA488CA6F1 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 360;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVDRQVAGW 10
Db 265 GHVKDQYTGW 274

RESULT 19
SNIP_MOUSE
ID -SNIP_MOUSE STANDARD; PRT; 1250 AA.
AC Q9QW16; O70298;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE p130Cas-associated protein (p140Cap) (SNAP-25-interacting protein)
DE (SNIP).
GN Name=p140; Synonyms=Kiaa1684;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS 1 AND 2).
RA Croci L., Bossolasco M., Gonzalez G.G.;
RL Submitted (DDB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 271-1250 (ISOFORM 3).
RC TISSUE=Brain;
RX PubMed=14621295;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Soga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT III. The complete nucleotide sequences of 500 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:167-180(2003).
RN [3]
RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX PubMed=14657239; DOI=10.1091/mbc.E03-09-0689;
RA Di Stefano P., Cabodi S., Erba E.B., Margaria V., Bergatto E.,
RA Guiffrida M.G., Silengo L., Tarone G., Turco E., Defilippi P.;
RT "p130Cas-associated protein (p140Cap) as a new tyrosine-phosphorylated
RT protein involved in cell spreading.";
RL Mol. Biol. Cell 15:787-800(2004).
RN [4]
RP PHOSPHORYLATION SITE SER-1054.
RX PubMed=15572359; DOI=10.1074/jbc.M411220200;
RA Collins M.O., Yu L., Coda M.P., Huel H., Campuzano I.,

Blackstock W.P., Choudhary J.S., Grant S.G.;
"Proteomic analysis of in vivo phosphorylated synaptic proteins.";
J. Biol. Chem. 280:5972-5982(2005).
-1- FUNCTION: Delays the onset of cell spreading in the early stages
of cell adhesion to fibronectin. Also involved in calcium-
dependent exocytosis from PC12 cells (By similarity).
-1- SUBUNIT: Interacts with BCAR1/p130Cas through its C-terminal
domain. Also interacts with the N-terminal coiled-coil region of
SNAP25 (By similarity).
-1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=1; Synonyms=1a;
IsoId=Q9QW16-1; Sequence=Displayed;
Name=2; Synonyms=1b;
IsoId=Q9QW16-2; Sequence=VSP_050631;
Name=3;
IsoId=Q9QW16-3; Sequence=VSP_050632;
Note=No experimental confirmation available;
-1- TISSUE SPECIFICITY: Expressed mainly in brain, testis and
epithelial-rich tissues such as mammary gland, lung and kidney.
-1- PTM: Tyrosine-phosphorylated in response to EGF and to cell
adhesion to integrin ligands (By similarity).
-1- SIMILARITY: Belongs to the p140Cap family.
-1- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
in positions 483 and 527.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; AF040944; AAC15635.1; -; mRNA.
EMBL; U59873; AAD00087.1; -; mRNA.
EMBL; AK129422; BAC98232.1; ALT_FRAME; mRNA.
PIR; T34101; T34101.
DR Ensembl; ENSMUSG00000038453; Mus musculus.
DR MGI; MGI:1933179; P140.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007162; P:negative regulation of cell adhesion; IDA.
KW Alternative splicing; Coiled coil; Cytoskeleton; Exocytosis;
Phosphorylation.
FT REGION 681 731 Interaction with SNAP25 (By similarity).
FT COILED 630 654 Potential.
FT COILED 713 845 Potential.
FT COILED 994 1011 Potential.
FT MOD_RES 1054 1054 Phosphoserine.
FT VARSPPLIC 1 41 MQPQCLRRPALAWERTAEGRARSPREEVGPDRDPQGRGEP
- -> MGNAPSQ (in isoform 2).
/FTId=VSP_050631.
Missing (in isoform 3).
/FTId=VSP_050632.
NV -> KL (in Ref. 2).
P -> R (in Ref. 2).
DV -> EL (in Ref. 2).
M -> V (in Ref. 2).
C -> G (in Ref. 2).
FT CONFLICT 1121 1122 KL -> NV (in Ref. 1; AAD00087).
SQ SEQUENCE 1250 AA; 134859 MW; 85F394C9523F78C0 CRC64;

Query Match 69.6%; Score 39; DB 1; Length 1250;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 138 RMREQVGW 146

RESULT 20
Q4QDJ2_LEINA
ID Q4QDJ2_LEIMA PRELIMINARY; PRT; 1370 AA.
AC Q4QDJ2;
```

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=LmjF19.0290;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
 RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
 RA Saunders D., Seeger K., Warren T., Apostolou Z., Bauser C., Beck A.,
 RA Bianchetti G., Borzym K., Boche G., Bruschi C., Ciarloni L.,
 RA Duesterhoeft A., Fuchs M., Gabel C., Goffeau A., Hilbert H.,
 RA Klages S., Kube M., Marra M., Masuy D., Mueller-Auer S., Pohl T.,
 RA Purnelle B., Reinhardt R., Rieger M., Robben J., Schaefer M.,
 RA Tosato V., Volckaert G., Wambutt R., Wedler H., Zimmermann M.,
 RA Andreame M., and Barrell B.G.;/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RW EMBL; CTO05258; CAJ07114.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 1370 AA; 145591 MW; F96DB457E3230954 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 1370;
 Best Local Similarity 87.5%; Pred. No. 3.9e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VRDQVAGW 10
 Db 1110 VRQVAGW 1117

RESULT 21
 QSLW4_SILPO PRELIMINARY; PRT; 1808 AA.
 ID QSLW4_SILPO PRELIMINARY;
 AC QSLW4;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE PAN domain protein.
 GN OrderedLocNames=SP00133;
 OS Silicibacter pomeroyi.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 OC Rhodobacteraceae; Silicibacter.
 OX NCBI_TaxID=89184;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
 RX PubMed=15602564; DOI=10.1038/nature03170;
 RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
 RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
 RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
 RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
 RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
 RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
 RA Haft D.H., Selengut J., Ward N.;
 RA "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
 RT marine environment."
 RL Nature 432:910-913(2004).
 DR EMBL; CP000031; AAV93461.1; -; Genomic DNA.
 DR GO; GO:0005576; C:extracellular region; IEA.
 DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0007596; P:blood coagulation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002890; A2M_N.
 DR InterPro; IPR011625; A2M_N.2.
 DR InterPro; IPR000177; Apple-
 DR InterPro; IPR003014; PAN.
 DR Pfam; PF01835; A2M_N.1.
 DR Pfam; PF07703; A2M_N.2; 1.

DR Pfam; PF00024; PAN; 1.
 DR SMART; SM00223; APPLE; 1.
 KW Complete proteome.
 SQ SEQUENCE 1808 AA; 192586 MW; A36D983174E69576 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 1808;
 Best Local Similarity 66.7%; Pred. No. 5.2e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
 Db 742 RAADQIAGW 750

RESULT 22
 Q8H1Z7_9ROSI PRELIMINARY; PRT; 167 AA.
 ID Q8H1Z7_9ROSI PRELIMINARY;
 AC Q8H1Z7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative citrus disease resistance protein Pt19 (Fragment).
 OS Citrus grandis x Poncirus trifoliata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Sapindales; Rutaceae; Citrus x Poncirus.
 OX NCBI_TaxID=206422;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Deng Z., Huang S., Ling P., Chen C., Yu C., Weber C.A., Moore G.A.,
 RA Gmitter F.G. Jr.;
 RT "Cloning and characterization of NBS-LRR class resistance-gene
 RT candidate sequences in citrus."
 RT Theor. Appl. Genet. 101:814-822(2000).
 RN [2]

RESULT 23
 Q62FH7_BURMA PRELIMINARY; PRT; 233 AA.
 ID Q62FH7_BURMA PRELIMINARY;
 AC Q62FH7;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocNames=BMA3057;
 OS Burkholderia mallei (Pseudomonas mallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 OX NCBI_TaxID=13373;
 RN [1]

Query Match 67.9%; Score 38; DB 2; Length 167;
 Best Local Similarity 44.4%; Pred. No. 73;
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
 Db 38 KIQDEIAGW 46

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ARCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidse T.D., DeBoy R.T., Dmitrov G.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RA "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU48462.1; -; Genomic_DNA.
DR TIGR; BMA3057; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 233 AA; 24465 MW; FF961AB6D98CF48C CRC64;

Query Match 67.9%; Score 38; DB 2; Length 233;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 201 RLRAQIAGW 209

RESULT 24
Q63VC8 BURPS
ID Q63VC8 BURPS PRELIMINARY; PRT; 234 AA.
AC Q63VC8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BPSL1314;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.B., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchaveit M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RA "Genomic plasticity of the causative agent of melioidosis,
RA Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH35311.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 234 AA; 24596 MW; 8237542235E1815B CRC64;

Query Match 67.9%; Score 38; DB 2; Length 234;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
Db 202 RLRAQIAGW 210

RESULT 25
Q4WRF3 ASPFU
ID Q4WRF3 ASPFU PRELIMINARY; PRT; 243 AA.
AC Q4WRF3;

```

```

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Heme/steroid binding domain protein, putative.
GN ORFNames=Afulg16510;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.-Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Gobie A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Litage J.-P., Li W., Lord A., Lu C.,
RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penlva M.A., Pertea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RA "Genomic sequence of the pathogenic and allergenic filamentous fungus
RA Aspergillus fumigatus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHF01000004; EAL90979.1; -; Genomic_DNA.
SQ SEQUENCE 243 AA; 27517 MW; 0DF9E6357B6B37D1 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 243;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVDQVAGW 10
Db 189 GKVRKQVAHW 198

RESULT 26
DHAA_MYCBO
ID DHAA_MYCBO STANDARD; PRT; 300 AA.
AC Q9XB14; Q6EUU9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Haloalkane dehalogenase (EC 3.8.1.5).
GN Name=dhaa; Synonyms=dmbA; OrderedLocusNames=Mb2610;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=5033/66, and MULL;
RA Jesenska A., Strouhal M., Pavlova M., Tesinska I., Monincova M.,
RA Bartos M., Pavlik I., Rychlik I., Nagata Y., Damborsky J.;
RA "Mycobacterial haloalkane dehalogenases: cloning, biochemical
RA properties and distribution.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP

```

```

RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=CDC 1551 / OshKosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RX Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Emdolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- FUNCTION: Catalyzes hydrolytic cleavage of carbon-halogen bonds in
CC halogenated aliphatic compounds, leading to the formation of the
CC corresponding primary alcohols, halide ions and protons (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: 1-haloalkane + H(2)O = a primary alcohol +
CC halide.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: Belongs to the haloalkane dehalogenase family. Type 2
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BX842580; CA555500.1; -; Genomic DNA.
DR EMBL; AE000516; AAK46969.1; -; Genomic DNA.
DR PIR; B70725; B70725.
DR HSSP; P51698; 11Z7.
DR TIGR; MT2656; -.
DR TubercuList; RV2579; -.
DR HAWAP; MF_01231; -; 1.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000639; Epox hydrolase.
DR InterPro; IPR000379; Ser setrs.
DR Pfam; PF00561; Abhydrolase 1; 1.
DR PRINTS; PR00412; EPOXHYDRLASE.
DR Complete proteome; Hydrolase.
KW ACT_SITE 109 109 Nucleophile (By similarity).
FT ACT_SITE 133 133 Proton donor (By similarity).
FT ACT_SITE 273 273 Proton acceptor (By similarity).
SQ SEQUENCE 300 AA; 33728 MW; B8F4854749F22562 CRC64;
Query Match 67.9%; Score 38; DB 1; Length 300;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GRVRDQVAGW 10
Db ||:|||||
Db 252 GRIDYVRSW 261
RESULT 28
Q7UV53 RHOBAC PRELIMINARY; PRT; 349 AA.
AC Q7UV53;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Ordered locus names=RB2871;
OS Rhodopirellula baltica.

```

```

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schleutner H., Anann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294137; CAD72873.1; -; Genomic_DNA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR010372; DNA_pol3_delta.
DR InterPro; IPR005790; DNA_polIII_delta.
DR Pfam; PF06144; DNA_pol3_delta; I.
DR TIGRPFAMs; TIGR01128; hofA; I.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 349 AA; 38362 MW; 5AC4590EB5845D82 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 349;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VRDQVAGW 10
Db 206 VRDQVAGW 213
|||||

RESULT 29
Q4H705_9DEIO
ID Q4H705_9DEIO PRELIMINARY; PRT; 367 AA.
AC Q4H705;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Peptidase M20:Peptidase M20.
GN ORFNames=DgeODRAFT_0250;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Ierani S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000011; EAL82057.1; -; Genomic DNA.
SQ SEQUENCE 367 AA; 39038 MW; 84B167B157FC3CF8 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 367;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVDQVAGW 10

```

```

Db 32 GRVMDVVEGW 41
|||||

RESULT 30
Q89SF5_BRAJA
ID Q89SF5_BRAJA PRELIMINARY; PRT; 367 AA.
AC Q89SF5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN OrderedLocuNames=bl12445;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shingo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC47710.1; -; Genomic_DNA.
DR HSP; Q51990; LGWJ.
DR GO; GO:0016491; F:Oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001155; Oxidored_FMN.
DR Pfam; PF00724; Oxidored_FMN; I.
KW Complete proteome.
SQ SEQUENCE 367 AA; 39432 MW; 791C46C4EE39BE4C CRC64;

Query Match 67.9%; Score 38; DB 2; Length 367;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RDQVAGW 10
Db 79 KDQVAGW 85
|||||

RESULT 31
Q747F0_GEOSL
ID Q747F0_GEOSL PRELIMINARY; PRT; 369 AA.
AC Q747F0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=GSU3317;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969(2003).

```



```
DR EMBL; AE017180; AAR36707.1; -; Genomic_DNA.
DR TIGR; GSU3317; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 369 AA; 37161 MW; 1A22761D0A1DAA7D CRC64;

Query Match 67.9%; Score 38; DB 2; Length 369;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
   |:|:|
Db 80 GKIRGPIAGW 89

RESULT 32
Q9L144_STRCO
ID Q9L144_STRCO PRELIMINARY; PRT; 385 AA.
AC Q9L144
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative racemase.
GN OrderedLocusNames=SCO7415; ORFNames=SC6D11.11;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiese T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939131; CAB76334.1; -; Genomic_DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001354; MR_MLE.
DR Pfam; PF01188; MR_MLE; 1.
DR Pfam; PF02746; MR_MLE_N; 1.
KW Complete proteome.
SQ SEQUENCE 385 AA; 41747 MW; 9777B77B4A2EB0B7 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 385;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
   |:|:|
Db 159 RLTDQLAGW 167

RESULT 33
Q74SA4_YERPE
ID Q74SA4_YERPE PRELIMINARY; PRT; 443 AA.
AC Q74SA4
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Signal transduction histidine kinase.
GN Namesbaes8; OrderedLocusNames=Y2718;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
```

```
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans";
RL DNA Res. 11:179-197(2004).
DR EMBL; AE017137; AAS62907.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008358; Lantibiot_regn.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR PRINTS; PR01780; LANTIREGPROT.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR PROSITE; PS50985; HAMP; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
KW Kinase.
SQ SEQUENCE 443 AA; 49971 MW; 026FA684A5751F4B CRC64;

Query Match 67.9%; Score 38; DB 2; Length 443;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
   |:|:|
Db 113 KVNDQVVGW 121

RESULT 34
Q66HY4_BRARE
ID Q66HY4_BRARE PRELIMINARY; PRT; 448 AA.
AC Q66HY4
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:92214.
GN ORFNames=zgc:92214;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
```

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettunen M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kryzinski M.T., Skalska U., Smalhus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Director MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL, BC081623; AAH81623.1; -, mRNA.
DR Ensembl, ENSDARG00000036491; Danio rerio.
DR ZFIN, ZDB-GENE-040912-153; zgc:92214.
DR GO, GO:0016020; C:membrane; IEA.
DR GO, GO:0046872; F:metal ion binding; IEA.
DR GO, GO:0004497; F:monooxygenase activity; IEA.
DR GO, GO:0006118; P:electron transport; IEA.
DR InterPro, IPR001128; Cytochrome_P450.
DR InterPro, IPR002403; EP450IV.
DR Pfam, PF00067; P450; 1.
DR PRINTS, PR00465; EP450IV.
DR PRINTS, PR00385; P450.
DR PROSITE, PS00086; CYTOCHROME_P450; 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 448 AA; 51781 MW; 342AB9E9856CBDB5 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 448;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|||:|:|:
DB 193 RVKDNISGW 201

RESULT 35
Q668C3_YERPS
ID Q668C3_YERPS PRELIMINARY; PRT; 461 AA.
AC Q668C3
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Sensor protein baes (EC 2.7.3.-).
GN Name=baes; OrderedLocusNames=YPTB2817;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype 1;
EX PubMed-15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
comparison with Yersinia pseudotuberculosis";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831 (2004).
DR EMBL, BX936398; CAH22055.1; -, Genomic_DNA.
DR GO, GO:0016020; C:membrane; IEA.
DR GO, GO:0005524; F:ATP binding; IEA.
DR GO, GO:0016301; F:kinase activity; IEA.
DR GO, GO:0000155; F:two-component sensor molecule activity; IEA.

DR GO, GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO, GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro, IPR003594; ATPbind_Alpase.
DR InterPro, IPR005467; His_kinase.
DR InterPro, IPR003661; His_kinA_N.
DR InterPro, IPR003660; His_kin_HAMP.
DR InterPro, IPR004358; His_kin_like_C.
DR InterPro, IPR008358; Lantibiot_regn.
DR Pfam, PF00672; HAMP; 1.
DR Pfam, PF02519; HATPase_c; 1.
DR Pfam, PF00512; Hiska; 1.
DR PRINTS, PR00344; BCTRLSENSOR.
DR PRINTS, PR01780; LANTIREGPROT.
DR SMART, SM00304; HAMP; 1.
DR SMART, SM00387; HATPase_c; 1.
DR SMART, SM00388; Hiska; 1.
DR PROSITE, PS00885; HAMP; 1.
DR PROSITE, PS0109; HIS_KIN; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 461 AA; 51937 MW; 941A4EEEL177AD4F4 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 461;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
|||:|:|:
DB 131 KVNDQVVGW 139

RESULT 36
Q7X541_9ACTO
ID Q7X541_9ACTO PRELIMINARY; PRT; 588 AA.
AC Q7X541;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Acyl-CoA synthase.
GN Name=lipA;
OS Actinoplanes friuliensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micromonosporineae; Micromonosporaceae; Actinoplanes.
OX NCBI_TaxID=196914;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heinzelmann E., Berger S., Muller C., Hartner T., Poralla K.,
RA Wohleben W., Schwartz D.;
RT "An acyl-CoA dehydrogenase is involved in the formation of the cis3
doublebond in the acyl residue of the lipopeptide antibiotic
frutitimicin in Actinoplanes friuliensis.";
RL Microbiol. 151:1963-1974 (2005).
RL EMBL, AJ488769; CAD32910.2; -, Genomic_DNA.
DR InterPro, IPR000873; AMP-bind.
DR Pfam, PF00501; AMP-binding; 1.
SQ SEQUENCE 588 AA; 62629 MW; AB80EBC26969D107 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 588;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
|||:|:|:
DB 124 GESRDEVVGW 133

RESULT 37
Q869Y0_DICDI
ID Q869Y0_DICDI PRELIMINARY; PRT; 941 AA.
AC Q869Y0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to plasmodium falciparum. cell differentiation protein rccl1,

Mon May 15 11:35:33 2006

```
DE putative.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RX MEDLINE=22097622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J.A.,
RA Bankier A.F., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A.,
RA Platzner M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Baumgart C.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC116305; AA052297.1; -; Genomic_DNA.
DR InterPro; IPR007216; Rcd1.
DR PANTHER; PTHR12262; Rcd1; 1.
DR Pfam; PF04078; Rcd1; 1.
SQ SEQUENCE 941 AA; 107408 MW; E01F97A1CFE717CA7 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 941;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
DB 740 KIRDQLIGW 748
:::|::|

RESULT 38
Q553Y8 D1CDDI PRELIMINARY; PRT; 941 AA.
AC Q553Y8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cell differentiation family, Rcd1-like protein.
GN Name=DDB0220000; ORFNames=DDB0220000;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Loulleged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
RA Chieholm R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AA01000026; EAL69925.1; -; Genomic_DNA.
DR
```

```
DR InterPro; IPR007216; Rcd1.
DR Pfam; PF04078; Rcd1; 1.
DR Nucleotide-binding.
SQ SEQUENCE 941 AA; 107382 MW; 5D36FB71CFE71079 CRC64;

Query Match 67.9%; Score 38; DB 2; Length 941;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 RVRDQVAGW 10
DB 740 KIRDQLIGW 748
:::|::|

RESULT 39
Q8UE69 AGRT5 PRELIMINARY; PRT; 63 AA.
AC Q8UE69;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Atul1892.
GN OrderedLocustNames=Atul1892;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A., Gordon D.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Krespan W., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
DR EMBL; AE009143; AAL42888.1; -; Genomic_DNA.
DR PIR; AB2809; AB2809.
DR InterPro; IPR010710; DUF1289.
DR Pfam; PF06945; DUF1289; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 63 AA; 7083 MW; C614BB7BEDC5F4B4 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 63;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GRVRDQVAGW 10
DB 22 GRTLDEIGW 31
:::|::|

RESULT 40
Q4LUS0 9BURK PRELIMINARY; PRT; 121 AA.
ID Q4LUS0_9BURK PRELIMINARY;
AC Q4LUS0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Response regulator receiver.
GN ORFNames=Bcen2424DRAFT_4960;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
```

```

RN  NUCLEOTIDE SEQUENCE.
RP  STRAIN=H12424;
RC  US DOE Joint Genome Institute (JGI-PGF);
RG  Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA  Hammon N., Ierani S., Pitluck S., Richardson P.;
RT  "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RL  H12424."
RL  Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN  NUCLEOTIDE SEQUENCE.
RP  STRAIN=H12424;
RC  US DOE Joint Genome Institute (JGI-ORNL);
RG  Larimer F., Land M.;
RA  "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT  H12424."
RL  Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC  -!- CAUTION: The sequence shown here is derived from an
CC  EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC  preliminary data.
DR  EMBL; AAH01000013; EAM19906.1; -; Genomic DNA.
SQ  SEQUENCE 121 AA; 12687 MW; 5E592333A7F898D4 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 121;
Best Local Similarity 70.0%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
    |||||:
Db 11 GTVRDEVAGF 20

RESULT 41
ID Q606W1 METCA PRELIMINARY; PRT; 178 AA.
AC Q606W1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Conserved domain protein.
GN OrderedLocusNames=MCA1902;
OS Methylococcus capsulatus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylcoccales;
OC Methylococcaceae; Methylococcus.
OX NCBI_TaxID=414;
RN  NUCLEOTIDE SEQUENCE.
RP  STRAIN=Bath / NCIMB 11132;
RC PubMed=15383840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.L., Larsen O., Sakwa J., Bruseeth L., Khouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Methe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren Q., Read T.D., DeBoy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindhaug S.H., Holt I.E., Eidhammer I., Jonassen I., Vanaken S.,
RA Utterback T.R., Feldblyum T.V., Fraser C.M., Lilienhaug J.R.,
RA Eisen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).";
RL PLoS Biol. 2:1616-1628(2004).
DR EMBL; AE017282; AAU91823.1; -; Genomic_DNA.
DR TIGR; MCA1902; -.
KW Complete proteome.
SQ SEQUENCE 178 AA; 19998 MW; 10916536F02CABC7 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 178;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
    |||||:
Db 111 GAVADEAAGW 120

```

```

RESULT 42
Q9HRC7 HALSA PRELIMINARY; PRT; 181 AA.
AC Q9HRC7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription initiation factor IIE alpha subunit.
GN Name=tfeI; OrderedLocusNames=VNG0757G;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN  NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP  STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RC MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahaitas G.G., Berquist B., Pan M.,
RA Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madlocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlshocher M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005019; AAG19231.1; -; Genomic_DNA.
DR PIR; C84233; C84233.
DR GO; GO:0005673; C:transcription factor TFII complex; IEA.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006367; P:transcription initiation from RNA polymeras. .; IEA.
DR InterPro; IPR02853; TFII alpha.
DR Pfam; PF02002; TFII alpha; 1.
KW Complete proteome; Initiation factor.
SQ SEQUENCE 181 AA; 20937 MW; 06185D1B43393B5D CRC64;

Query Match 66.1%; Score 37; DB 2; Length 181;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
    |||||:
Db 78 RVRDEDSGW 86

RESULT 43
ALBR_KLEOX
ID ALBR_KLEOX STANDARD; PRT; 218 AA.
AC P10488;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Albicidin resistance protein.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
RN  NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP  STRAIN=JMP4505;
RC MEDLINE=89013885; PubMed=2845223;
RA Walker M.J., Birch R.G., Pemberton J.M.;
RT "Cloning and characterization of an albicidin resistance gene from
RT Klebsiella oxytoca."
RL Mol. Microbiol. 2:443-454(1988).
CC -!- FUNCTION: Albicidin resistance protein binds to form a complex
CC without antibiotic activity but without catalyzing any further
CC chemical modifications to albicidin.
CC -!- SUBCELLULAR LOCATION: Periplasmic.

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; Y00558; CAA68640.1; -; Genomic_DNA.

DR PIR; S02828; S02828; Albicidin.res.

DR InterPro; IPR008396; Albicidin.res.

DR Pfam; PF05583; Albicidin.res; 2.

KW Antibiotic resistance; Periplasmic.

SQ SEQUENCE 218 AA; 25859 MW; 56578223BBDCEBA7 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 218;

Best Local Similarity 70.0%; Pred. No. 1.5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 1;

QY 1 GRVRDQVAGW 10

||:|:|:|:|

48 GLRLRWQAGW 57

RESULT 44

Q67J54 SYMTH

ID Q67J54_SYMTH PRELIMINARY; PRT; 244 AA.

AC Q67J54;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Formate dehydrogenase associated protein.

GN OrderedLocNames=STH3094;

OS Symbiobacterium thermophilum.

OC Bacteria; Actinobacteria; Symbiobacterium.

OX NCBI_TaxID=2734;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=IAM14863;

RX PubMed=15383646; DOI=10.1093/nar/gkh830;

RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsui T.,

RA Morimura K., Ikeda H., Hattori M., Beppu T.;

RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable

RT bacterium that depends on microbial commensalism.";

RL Nucleic Acids Res. 32:4937-4944(2004).

DR EMBL; AF008840; BAD42076.1; -; Genomic_DNA.

DR GO; GO:0009326; C:formate dehydrogenase complex; IEA.

DR GO; GO:0008863; F:formate dehydrogenase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR003786; FdhD.

DR InterPro; IPR000408; Reg_chr_condens.

DR Pfam; PF02634; FdhD-NarQ; 1.

DR PROSITE; PS00626; RCC1_2; 1.

KW Complete proteome.

SQ SEQUENCE 244 AA; 25720 MW; F5566303CAE2ABD6 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 244;

Best Local Similarity 85.7%; Pred. No. 1.6e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 4 RDQVAGW 10

||:|:|:|:|

114 REQVAGW 120

RESULT 45

Q5UQ11 MIMIV

ID Q5UQ11_MIMIV PRELIMINARY; PRT; 276 AA.

AC Q5UQ11;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)

DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Hypothetical protein.

GN ORFNames=MIMI_L191;

OS Mimivirus.

OC Viruses; dsDNA viruses, no RNA stage; Mimivirus.

OX NCBI_TaxID=212035;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Rowbotham-Bradford;

RX MEDLINE=22550848; PubMed=12663918; DOI=10.1126/science.1081867;

RA La Scola B., Audic S., Robert C., Jungang L., de Lamballerie X.,

RA Drancourt M., Birtles R., Claverie J.M., Raoult D.;

RT "A giant virus in amoebae.";

RL Science 299:2033-2033(2003).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Rowbotham-Bradford;

RX PubMed=15486256; DOI=10.1126/science.1101485;

RA Raoult D., Audic S., Robert C., Abergel C., Renesto P., Ogata H.,

RA La Scola B., Susan M., Claverie J.M.;

RT "The 1.2-Mb Genome Sequence of Mimivirus.";

RL Science 306:1344-1350(2004).

DR EMBL; AY653733; AAV50465.1; -; Genomic_DNA.

KW Hypothetical protein.

SQ SEQUENCE 276 AA; 32656 MW; 6EE1DBAC6433055C CRC64;

Query Match 66.1%; Score 37; DB 2; Length 276;

Best Local Similarity 66.7%; Pred. No. 1.9e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 6; Conservative 1;

QY 2 RVRDQVAGW 10

||:|:|:|:|

28 RVQDTIGW 36

RESULT 46

Q4UPY2 XANCP

ID Q4UPY2_XANCP PRELIMINARY; PRT; 307 AA.

AC Q4UPY2;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=XC_3851;

OS Xanthomonas campestris pv. campestris str. 8004.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI_TaxID=314565;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=8004;

RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,

RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,

RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,

RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,

RA He C.-Z.;

RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; CP000050; AAY50891.1; -; Genomic_DNA.

KW Hypothetical protein.

SQ SEQUENCE 307 AA; 33115 MW; 141D08229A44B82A CRC64;

Query Match 66.1%; Score 37; DB 2; Length 307;

Best Local Similarity 55.6%; Pred. No. 2.1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 5; Conservative 2;

QY 2 RVRDQVAGW 10

||:|:|:|:|

85 RIRQSIAGW 93

RESULT 47

Q4NPES_9DEL

ID Q4NPES_9DEL PRELIMINARY; PRT; 307 AA.

AC Q4NPES;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Putative oxidoreductase.
GN ORFNames=AdelDRAFT_0539;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytochromes; Myxococcales; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hannon N., Irani S., Pitluck S., Richardson P.,
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.,
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000045; EAL77424.1; -; Genomic DNA.
SQ SEQUENCE 307 AA; 32136 MW; CA8F0919745F797F CRC64;

Query Match 66.1%; Score 37; DB 2; Length 307;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GRVRDQVAGW 10
|:|:|:|:
Db 296 GLIQDVVAGW 305

RESULT 48
Q8P4D1_XANCP PRELIMINARY; PRT; 307 AA.
AC Q8P4D1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XCC3779.
GN OrderedLocNames=XCC3779;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Anaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Clapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teal S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).

DR EMBL; AE012498; AAM43479.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 307 AA; 33114 MW; 141D08229A44B82A CRC64;

Query Match 66.1%; Score 37; DB 2; Length 307;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|:|:|:
Db 85 RIRQSIAGW 93

RESULT 49
Q4H433_SOLITU PRELIMINARY; PRT; 335 AA.
AC Q4H433;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hsr203J.
GN Solanum tuberosum (potato).
OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC asterids; lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yamamoto C., Yoshioka H.;
RT "Hypersensitive response marker."
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB200918; BAE16319.1; -; mRNA.
SQ SEQUENCE 335 AA; 37308 MW; 8D00B1F7A2F61D19 CRC64;

Query Match 66.1%; Score 37; DB 2; Length 335;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RVRDQVAGW 10
|:|:|:|:
Db 6 RVKDVVSGW 14

RESULT 50
Q50N02_ENTHI PRELIMINARY; PRT; 372 AA.
AC Q50N02;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=470.t00006;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suhr B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoef U., Bhattacharya A.,
RA Chillingworth T., Church C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitsch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica."
RL Nature 433:865-868(2005).

CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFH01001129; EAL42974.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 372 AA; 42436 MW; F5B8824993DD752D CRC64;

Query Match 66.1%; Score 37; DB 2; Length 372;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GRVRDQVAGW 10
|:|:|:|:
Db 111 GKVRDLIATW 120

Search completed: May 12, 2006, 10:48:54
Job time : 120.769 secs

THIS PAGE BLANK (USPTO)

98	33	67.3	490	5	ABB55083	Abb55083 Lactococc	171	31	63.3	15	2	AAy22380	Aay22380 TPO recep
99	33	67.3	673	6	ABU33383	Abu33383 Protein e	172	31	63.3	55	4	AAm21674	Aam21674 Peptide #
100	33	67.3	687	9	ABE39948	Abe39948 L. pneumo	173	31	63.3	55	4	AAm37981	Aam37981 Peptide #
101	33	67.3	687	9	ABE36554	Abe36554 L. pneumo	174	31	63.3	55	4	AAm65049	Aam65049 Human bra
102	33	67.3	780	7	ABM88746	Abm88746 Rice abio	175	31	63.3	55	5	ABg46787	Abg46787 Human pep
103	33	67.3	793	5	ABJ17937	Abj17937 Rabbit ne	176	31	63.3	109	4	ABG19372	Abg19372 Novel hum
104	33	67.3	862	2	AAV07064	Aav07064 Renal can	177	31	63.3	109	4	ABG11683	Abg11683 Novel hum
105	33	67.3	862	4	AAE62199	Aae62199 Amino aci	178	31	63.3	120	6	ABP80598	Abp80598 N. gonorr
106	33	67.3	862	4	ABR47576	Abr47576 Breast ca	179	31	63.3	120	6	ABU37678	Abu37678 Protein e
107	33	67.3	862	7	ADJ70641	Adj70641 Human hea	180	31	63.3	120	6	ABU37281	Abu37281 Protein e
108	33	67.3	1153	6	ABU23672	Abu23672 Protein e	181	31	63.3	125	9	ABE38788	Abe38788 L. pneumo
109	33	67.3	1530	7	ABM88229	Abm88229 Rice abio	182	31	63.3	131	6	AAE34014	Aae34014 Lolium pe
110	33	67.3	1873	1	AAp95645	Aap95645 Rabbit se	183	31	63.3	152	3	AAE44889	Aae44889 Human sec
111	33	67.3	1873	2	AAr73055	Aar73055 Rabbit sk	184	31	63.3	161	2	AAy35809	Aay35809 Amino aci
112	33	67.3	1873	2	AAW18390	Aaw18390 Rabbit ca	185	31	63.3	161	5	ABP28893	Abp28893 Streptoco
113	33	67.3	1873	2	AAW37711	Aaw37711 Rabbit sk	186	31	63.3	165	5	ABP28646	Abp28646 Streptoco
114	33	67.3	1873	3	AAy77544	Aay77544 Rabbit sk	187	31	63.3	165	8	ADV87871	Adv87871 Streptoco
115	33	67.3	1873	9	ABE26777	Abe26777 Rabbit sk	188	31	63.3	165	8	ADV81323	Adv81323 Streptoco
116	32	65.3	14	2	AAy22396	Aay22396 TPO recep	189	31	63.3	165	8	ADV79124	Adv79124 Streptoco
117	32	65.3	15	2	AAy22399	Aay22399 TPO recep	190	31	63.3	165	9	ABE41756	Abe41756 L. pneumo
118	32	65.3	15	2	AAy22381	Aay22381 TPO recep	191	31	63.3	175	9	ADX02798	Adx02798 Sendai vi
119	32	65.3	82	4	AAE63463	Aae63463 Human bre	192	31	63.3	176	9	ABE38502	Abe38502 L. pneumo
120	32	65.3	125	5	AAE16594	Aae16594 Human 576	193	31	63.3	177	9	ABE38841	Abe38841 L. pneumo
121	32	65.3	134	8	ADS18428	Ads18428 Entamoeba	194	31	63.3	181	9	ADX02797	Adx02797 Sendai vi
122	32	65.3	166	6	ABM69803	Abm69803 Photorhab	195	31	63.3	184	7	ADf03934	Adf03934 Bacterial
123	32	65.3	241	8	ADN26512	Adn26512 Bacterial	196	31	63.3	203	1	AAp70394	Aap70394 Sequence
124	32	65.3	251	8	ADQ36875	Adq36875 Cell prol	197	31	63.3	204	7	ADP0394	Adp0394 Sendai vi
125	32	65.3	275	6	ABU43401	Abu43401 Protein e	198	31	63.3	204	9	ADX02796	Adx02796 Sendai vi
126	32	65.3	315	7	ABM73684	Abm73684 DNA clone	199	31	63.3	206	6	ABU02574	Abu02574 S. pneumo
127	32	65.3	330	7	ABO65232	AbO65232 Klebsiell	200	31	63.3	206	8	ADK48267	Adk48267 Streptoco
128	32	65.3	349	4	AAE05984	Aae05984 Lactobaci	201	31	63.3	207	8	ADR94944	Adr94944 Novel S.
129	32	65.3	349	7	ADH89124	Adh89124 L. rhamno	202	31	63.3	207	9	AAE58814	Aae58814 Streptoco
130	32	65.3	376	8	ADN19201	Adn19201 Bacterial	203	31	63.3	214	7	ABO83235	AbO83235 Pseudomon
131	32	65.3	383	6	ABU25688	Abu25688 Protein e	204	31	63.3	215	8	ADX78070	Adx78070 Plant ful
132	32	65.3	398	7	ABM85352	Abm85352 Mouse pro	205	31	63.3	215	9	ADX02795	Adx02795 Sendai vi
133	32	65.3	402	8	ADS26395	Ads26395 Bacterial	206	31	63.3	226	8	ADK48371	Adk48371 Streptoco
134	32	65.3	403	8	ADS28482	Ads28482 Bacterial	207	31	63.3	246	8	ADR94895	Adr94895 Novel S.
135	32	65.3	403	8	ADN26837	Adn26837 Bacterial	208	31	63.3	246	9	AAE58765	Aae58765 Streptoco
136	32	65.3	403	8	ADS26763	Ads26763 Bacterial	209	31	63.3	258	6	ADA35565	Ada35565 Acinetoba
137	32	65.3	403	8	ADS27145	Ads27145 Bacterial	210	31	63.3	272	4	ABG11685	Abg11685 Novel hum
138	32	65.3	407	8	ADP84838	Adp84838 Cellobioh	211	31	63.3	279	3	AAE29865	Aae29865 Human sec
139	32	65.3	410	8	ADS44448	Ads44448 Bacterial	212	31	63.3	283	3	AAg13091	Aag13091 Arabidops
140	32	65.3	535	4	ABG15415	Abg15415 Novel hum	213	31	63.3	284	3	AAg13090	Aag13090 Arabidops
141	32	65.3	793	2	AAW25635	Aaw25635 Human cad	214	31	63.3	287	8	ADS22304	Ads22304 Bacterial
142	32	65.3	793	2	AAW13132	Aaw13132 Full leng	215	31	63.3	291	5	ABG13089	Abg13089 Arabidops
143	32	65.3	799	2	AAW25639	Aaw25639 Rat full	216	31	63.3	291	5	ABG91527	Abg91527 Purine/py
144	32	65.3	799	2	AAW13126	Aaw13126 Full leng	217	31	63.3	291	8	ADI43725	Adi43725 Plant tra
145	32	65.3	799	8	ADG42065	Adg42065 Human CDH	218	31	63.3	291	8	ADS43208	Ads43208 Bacterial
146	32	65.3	799	8	ADP26902	Adp26902 Human cad	219	31	63.3	291	9	AEA26845	Aea26845 Stress to
147	32	65.3	810	6	ABM70687	Abm70687 Photorhab	220	31	63.3	292	5	ABB48277	Abb48277 Listeria
148	32	65.3	824	4	ABE52696	Abe52696 Escherich	221	31	63.3	293	4	AAE96312	Aae96312 Putative
149	32	65.3	832	4	AAE84218	Aae84218 Amino aci	222	31	63.3	315	7	ADE57869	Ade57869 Human Pro
150	32	65.3	845	4	AAE80134	Aae80134 Human pro	223	31	63.3	315	7	ADE57877	Ade57877 Human Pro
151	32	65.3	847	6	ABM70680	Abm70680 Photorhab	224	31	63.3	315	7	ADE57885	Ade57885 Human Pro
152	32	65.3	848	7	ABO67256	AbO67256 Klebsiell	225	31	63.3	315	7	ADE57881	Ade57881 Human Pro
153	32	65.3	852	6	ABM70692	Abm70692 Photorhab	226	31	63.3	315	7	ADE57889	Ade57889 Human Pro
154	32	65.3	854	4	AAE84217	Aae84217 Amino aci	227	31	63.3	315	7	ADE57873	Ade57873 Human Pro
155	32	65.3	854	4	AAW79150	Aaw79150 Human pro	228	31	63.3	315	8	ADP90915	Adp90915 Human ald
156	32	65.3	854	5	ABE81932	Abe81932 Human Dis	229	31	63.3	315	9	ADV85532	Adv85532 Human ald
157	32	65.3	854	8	ADK17424	Adk17424 Human Dis	230	31	63.3	316	2	AAE15425	Aae15425 Placenta-
158	32	65.3	854	8	ADR47379	Adr47379 Human Dis	231	31	63.3	316	2	AAE15425	Aae15425 Human ald
159	32	65.3	863	8	ADR47391	Adr47391 DISC1-FA.	232	31	63.3	316	5	ABB04753	Abb04753 Human AKR
160	32	65.3	896	4	AAE00820	Aae00820 Human bon	233	31	63.3	316	5	ABB77443	Abb77443 Human tum
161	32	65.3	940	6	ABU49109	Abu49109 Protein e	234	31	63.3	316	7	ADJ69767	Adj69767 Human hea
162	32	65.3	1195	8	ADS21995	Ads21995 Bacterial	235	31	63.3	316	8	ADQ15140	Adq15140 Human can
163	32	65.3	1231	7	ADJ68516	Adj68516 Human hea	236	31	63.3	316	8	ABM80940	Abm80940 Tumour-as
164	32	65.3	1246	8	ADQ97435	Adq97435 Human can	237	31	63.3	316	8	ADP23164	Adp23164 PRO polyP
165	32	65.3	1246	8	ABM80794	Abm80794 Tumour-as	238	31	63.3	316	9	ABE22075	Abe22075 Human AKR
166	32	65.3	1266	8	ADQ51507	Adq51507 Human jum	239	31	63.3	318	5	AAE15443	Aae15443 Human dru
167	32	65.3	1456	7	ADM04577	Adm04577 Human pro	240	31	63.3	326	9	ABE42007	Abe42007 L. pneumo
168	32	65.3	1620	6	ABP98333	Abp98333 Amino aci	241	31	63.3	328	4	ABG11687	Abg11687 Novel hum
169	32	65.3	2434	3	AAE42264	Aae42264 Human ORF	242	31	63.3	347	9	ABE42011	Abe42011 L. pneumo
170	32	65.3	2594	4	AAE005984	Aae005984 Human bon	243	31	63.3	348	6	ABU17065	Abu17065 Protein e

244	31	63.3	356	6	ADA33310	AdA33310 Acinetoba	317	30	61.2	223	7	ADC10983	Adc10983 Human pro
245	31	63.3	391	9	AEb42009	Aeb42009 L. pneumo	318	30	61.2	232	9	AEb98137	Aeb98137 Protein e
246	31	63.3	391	9	AEb41974	Aeb41974 L. pneumo	319	30	61.2	238	4	AAG91726	Aag91726 C glutami
247	31	63.3	391	9	AEb41956	Aeb41956 L. pneumo	320	30	61.2	238	6	ABR44270	ABr44270 Novel hum
248	31	63.3	391	9	AEb42013	Aeb42013 L. pneumo	321	30	61.2	238	6	ABR44269	ABr44269 Novel hum
249	31	63.3	394	4	ABG19373	ABg19373 Novel hum	322	30	61.2	238	7	ADE14367	Adel14367 Human int
250	31	63.3	402	9	AEb38745	Aeb38745 L. pneumo	323	30	61.2	238	8	ABO84778	Abo84778 Human can
251	31	63.3	402	9	AEb38771	Aeb38771 L. pneumo	324	30	61.2	238	8	ADN25862	Adn25862 Bacterial
252	31	63.3	416	6	AEb34863	Aeb34863 Human kin	325	30	61.2	241	4	AAG98328	Aag98328 Escherich
253	31	63.3	416	8	ADK65919	Adk65919 Human IPM	326	30	61.2	241	4	ADS28386	AdS28386 Bacterial
254	31	63.3	433	8	ADK16215	Adk16215 Nanoarcha	327	30	61.2	251	8	ADR95327	Adr95327 Novel S.
255	31	63.3	437	6	AEb33339	Aeb33339 Protein e	328	30	61.2	251	8	AEb59197	Aeb59197 Streptoco
256	31	63.3	437	6	AEb33769	Aeb33769 L. pneumo	329	30	61.2	254	3	AAG52497	Aag52497 Arabidops
257	31	63.3	437	9	AEb40152	Aeb40152 L. pneumo	330	30	61.2	255	3	AAG52496	Aag52496 Arabidops
258	31	63.3	450	6	ADb45062	Adb45062 Human pro	331	30	61.2	256	8	ABO59795	Abo59795 Human gen
259	31	63.3	451	5	ABF25826	ABf25826 Streptoco	332	30	61.2	285	2	ADN46138	Adn46138 Thermoco
260	31	63.3	451	8	ADV89528	Adv89528 Streptoco	333	30	61.2	285	2	AAV35437	Aay35437 Chlamydia
261	31	63.3	451	8	ADV80781	Adv80781 Streptoco	334	30	61.2	286	9	ABM96030	ABm96030 M. xanthu
262	31	63.3	451	8	ADV82942	Adv82942 Streptoco	335	30	61.2	291	7	ABM85353	ABm85353 Human pro
263	31	63.3	459	4	ABG11686	ABg11686 Novel hum	336	30	61.2	292	3	AAG52495	Aag52495 Arabidops
264	31	63.3	496	6	ABU33143	ABu33143 Protein e	337	30	61.2	302	4	AAb94000	AAb94000 Human pro
265	31	63.3	496	9	AEb40962	Aeb40962 L. pneumo	338	30	61.2	308	7	ABO79435	Abo79435 Pseudomon
266	31	63.3	496	9	AEb37651	Aeb37651 L. pneumo	339	30	61.2	308	7	ABO79435	Abo79435 Thermoco
267	31	63.3	637	8	ADX66365	Adx66365 Plant ful	340	30	61.2	315	8	ADb47480	Adb47480 Thermoco
268	31	63.3	644	6	ABU25152	ABu25152 Protein e	341	30	61.2	332	8	ADb65485	ADb65485 Klebsiell
269	31	63.3	687	2	AAb97123	AAb97123 Chicken d	342	30	61.2	336	9	ADZ25527	Adz25527 Cytokine
270	31	63.3	729	8	ADb27264	Adb27264 Bacterial	343	30	61.2	344	4	ABG26158	ABg26158 Novel hum
271	31	63.3	737	8	ADK47776	Adk47776 Streptoco	344	30	61.2	346	8	ADN22477	Adn22477 Bacterial
272	31	63.3	744	8	ADR95419	Adr95419 Novel S.	345	30	61.2	347	3	AAG51011	Aag51011 Arabidops
273	31	63.3	744	9	AEA59289	Aea59289 Streptoco	346	30	61.2	348	4	AAb76604	Aab76604 Corynebac
274	31	63.3	769	8	ADS27589	AdS27589 Bacterial	347	30	61.2	348	4	AAb79964	Aab79964 Corynebac
275	31	63.3	782	5	AAU98072	Aau98072 Chicken n	348	30	61.2	348	4	AAb79964	Aab79964 Corynebac
276	31	63.3	794	6	ABU17842	ABu17842 Protein e	349	30	61.2	348	4	AAb79964	Aab79964 Corynebac
277	31	63.3	842	4	ABG85019	ABg85019 Shrimp wh	350	30	61.2	350	4	AAO17160	AAo17160 Corynebac
278	31	63.3	863	4	ABB58669	ABb58669 Drosophil	351	30	61.2	350	4	AAb48049	AAb48049 Signal tr
279	31	63.3	924	5	ABB54116	ABb54116 Lactococ	352	30	61.2	351	4	AAb62149	AAb62149 P.falci
280	31	63.3	924	5	ADS29350	ADs29350 Bacterial	353	30	61.2	355	3	AAG28692	Aag28692 Arabidops
281	31	63.3	932	6	ABU19329	ABu19329 Protein e	354	30	61.2	359	4	AAb30608	AAb30608 Amino aci
282	31	63.3	978	8	ABM92566	ABm92566 Mouse can	355	30	61.2	359	5	ABP65172	ABp65172 Hypoxia-r
283	31	63.3	990	9	ABM92566	ABm92566 M. xanthu	356	30	61.2	359	5	ADP54847	ADp54847 Human pro
284	31	63.3	1016	8	ADR08525	Adr08525 Human pro	357	30	61.2	380	8	ADQ19952	Adq19952 Human sof
285	31	63.3	1164	8	ADN23911	Adn23911 Bacterial	358	30	61.2	380	8	ADQ19952	Adq19952 Human sof
286	31	63.3	1215	4	AAE02340	AAe02340 D. melano	359	30	61.2	382	6	ABU33633	ABu33633 Protein e
287	31	63.3	1264	8	ADQ67800	ADq67800 Novel hum	360	30	61.2	382	6	ABU33633	ABu33633 Protein e
288	31	63.3	2172	4	ABM67958	ABm67958 Drosophil	361	30	61.2	382	6	ABU33633	ABu33633 Protein e
289	31	63.3	2324	9	ABM90533	ABm90533 M. xanthu	362	30	61.2	385	4	ABG58316	ABg58316 Drosophil
290	30	61.2	11	2	AAV22367	Aay22367 TPO recep	363	30	61.2	392	3	AAG51010	Aag51010 Arabidops
291	30	61.2	15	2	AAV22367	Aay22367 TPO recep	364	30	61.2	394	6	ADb10152	ADb10152 Allolococ
292	30	61.2	15	5	ABP34908	ABp34908 Human ORF	365	30	61.2	399	3	AEb36128	Aeb36128 L. pneumo
293	30	61.2	62	5	ABP04777	ABp04777 Human ORF	366	30	61.2	400	3	AAG28691	Aag28691 Arabidops
294	30	61.2	93	9	AEb41682	Aeb41682 L. pneumo	367	30	61.2	402	4	AAb85362	AAb85362 Human pho
295	30	61.2	94	4	AAU01705	Aau01705 Propionib	368	30	61.2	404	5	ABP69581	ABp69581 Novel hum
296	30	61.2	99	9	AEb38411	Aeb38411 L. pneumo	369	30	61.2	411	5	ADQ65892	ADq65892 Lactococ
297	30	61.2	118	6	ABU50358	ABu50358 Protein e	370	30	61.2	422	5	ADQ65892	ADq65892 Lactococ
298	30	61.2	150	7	ADC89121	ADc89121 Ribosomal	371	30	61.2	432	6	ABM73527	ABm73527 Staphyloc
299	30	61.2	150	7	ADC89106	ADc89106 Ribosomal	372	30	61.2	434	3	AAm25227	AAm25227 Human pro
300	30	61.2	171	3	AAb58741	ABb58741 Breast an	373	30	61.2	446	8	ADX94435	ADx94435 Plant ful
301	30	61.2	171	5	ABP41060	ABp41060 Human ova	374	30	61.2	448	8	ADX94435	ADx94435 Plant ful
302	30	61.2	171	6	ABP41060	ABp41060 Human ova	375	30	61.2	453	3	AAG28690	Aag28690 Arabidops
303	30	61.2	176	4	AAU39170	Aau39170 Propionib	376	30	61.2	461	6	ABU19541	ABu19541 Protein e
304	30	61.2	176	5	ABP10173	ABp10173 Human ORF	377	30	61.2	466	4	ADN18586	Adn18586 Bacterial
305	30	61.2	176	6	ABM35689	ABm35689 Propionib	378	30	61.2	466	8	ADJ17050	Adj17050 Human hea
306	30	61.2	177	3	AAb51685	ABb51685 Gene 22 h	379	30	61.2	476	7	ADJ17050	Adj17050 Human hea
307	30	61.2	177	3	AAb51685	ABb51685 Human sec	380	30	61.2	487	5	AAU19641	Aau19641 Human nov
308	30	61.2	177	8	ABO59208	ABo59208 Human gen	381	30	61.2	487	5	AAU19641	Aau19641 Human nov
309	30	61.2	185	3	AAG58376	Aag58376 Arabidops	382	30	61.2	487	6	ABU17480	ABu17480 Protein e
310	30	61.2	202	9	AEb98133	Aeb98133 Protein e	383	30	61.2	487	6	ABU17480	ABu17480 Protein e
311	30	61.2	210	8	ADS24392	ADs24392 Bacterial	384	30	61.2	489	8	ADJ35156	Adj35156 C. tentan
312	30	61.2	210	8	AAG58375	Aag58375 Arabidops	385	30	61.2	517	8	ADJ35156	Adj35156 C. tentan
313	30	61.2	210	8	ADT56711	ADt56711 Plant pol	386	30	61.2	538	8	ADJ35156	Adj35156 C. tentan
314	30	61.2	212	9	AEb98131	Aeb98131 Protein e	387	30	61.2	546	2	AAW81975	AAw81975 Ehrlichia
315	30	61.2	223	4	AAU19801	Aau19801 Human nov	388	30	61.2	551	7	ABO67710	Abo67710 Klebsiell
316	30	61.2	223	5	ABP48021	ABp48021 Human pol	389	30	61.2	582	4	ABB69962	ABb69962 Drosophil

536	30	61.2	772	6	ABU91406	Abu91406 Human PRO	609	30	61.2	772	6	ABO27862	AbO27862 Human sec
537	30	61.2	772	6	ABU84620	Abu84620 Human sec	610	30	61.2	772	6	ABO29997	AbO29997 Human sec
538	30	61.2	772	6	ABR69710	AbR69710 Human sec	611	30	61.2	772	6	ABO33206	AbO33206 Human PRO
539	30	61.2	772	6	ABR69710	AbR69710 Human sec	612	30	61.2	772	6	ABM04894	Abm04894 Human sec
540	30	61.2	772	6	ABU93356	Abu93356 Human PRO	613	30	61.2	772	6	ABM08854	Abm08854 Human sec
541	30	61.2	772	6	ABO09909	AbO09909 Human sec	614	30	61.2	772	6	ABO36454	AbO36454 Human sec
542	30	61.2	772	6	ABO08994	AbO08994 Human sec	615	30	61.2	772	6	ABO35539	AbO35539 Human PRO
543	30	61.2	772	6	ABU10562	Abu10562 Human sec	616	30	61.2	772	6	ABO33504	AbO33504 Human sec
544	30	61.2	772	6	ABU95571	Abu95571 Human PRO	617	30	61.2	772	6	ABM10379	Abm10379 Human sec
545	30	61.2	772	6	ABU79809	Abu79809 Human sec	618	30	61.2	772	6	ABM11904	Abm11904 Human sec
546	30	61.2	772	6	ABU96780	Abu96780 Novel hum	619	30	61.2	772	6	ABO52050	AbO52050 Human PRO
547	30	61.2	772	6	ABR70625	AbR70625 Human sec	620	30	61.2	772	6	ABO52355	AbO52355 Human PRO
548	30	61.2	772	6	ABO04976	AbO04976 Novel hum	621	30	61.2	772	6	ABO23673	AbO23673 Human sec
549	30	61.2	772	6	ABO08384	AbO08384 Human sec	622	30	61.2	772	6	ABR97159	AbR97159 Human sec
550	30	61.2	772	6	ABO05591	AbO05591 Human sec	623	30	61.2	772	6	ABR86947	AbR86947 Human sec
551	30	61.2	772	6	ABR73980	AbR73980 Human sec	624	30	61.2	772	6	ABM10989	Abm10989 Human sec
552	30	61.2	772	6	ABR95572	AbR95572 Human sec	625	30	61.2	772	6	ABO28133	AbO28133 Human sec
553	30	61.2	772	6	ABR80869	AbR80869 Human sec	626	30	61.2	772	6	ABO32132	AbO32132 Human sec
554	30	61.2	772	6	ABR81174	AbR81174 Human sec	627	30	61.2	772	6	ABM15259	Abm15259 Human sec
555	30	61.2	772	6	ABN00870	Abn00870 Human sec	628	30	61.2	772	6	ABM06414	Abm06414 Human sec
556	30	61.2	772	6	ABR88472	AbR88472 Human sec	629	30	61.2	772	6	ABM04225	Abm04225 Human sec
557	30	61.2	772	6	ABM77293	Abm77293 Human sec	630	30	61.2	772	6	ABM22338	Abm22338 Human sec
558	30	61.2	772	6	ABO28777	AbO28777 Human sec	631	30	61.2	772	6	ABM07634	Abm07634 Human sec
559	30	61.2	772	6	ABO31522	AbO31522 Human sec	632	30	61.2	772	6	ABO40724	AbO40724 Human sec
560	30	61.2	772	6	ABM07939	Abm07939 Human sec	633	30	61.2	772	6	ABM35371	Abm35371 Human sec
561	30	61.2	772	6	ABO40419	AbO40419 Human sec	634	30	61.2	772	6	ABO33134	AbO33134 Human PRO
562	30	61.2	772	6	ABO35844	AbO35844 Human PRO	635	30	61.2	772	6	ABO52660	AbO52660 Human PRO
563	30	61.2	772	6	ADA77858	Ada77858 Human sec	636	30	61.2	772	6	ABO50220	AbO50220 Human sec
564	30	61.2	772	6	ABM24778	Abm24778 Human sec	637	30	61.2	772	6	ABU99214	Abu99214 Human sec
565	30	61.2	772	6	ABO03046	AbO03046 Human sec	638	30	61.2	772	6	ABO04266	AbO04266 Human sec
566	30	61.2	772	6	ABO03046	AbO03046 Human sec	639	30	61.2	772	6	ABO05896	AbO05896 Human sec
567	30	61.2	772	6	ABR90302	AbR90302 Human sec	640	30	61.2	772	6	ABM18436	Abm18436 Human sec
568	30	61.2	772	6	ABM17216	Abm17216 Human sec	641	30	61.2	772	6	ABR97464	AbR97464 Human sec
569	30	61.2	772	6	ABR95267	AbR95267 Human sec	642	30	61.2	772	6	ABM01175	Abm01175 Human sec
570	30	61.2	772	6	ABR95267	AbR95267 Human sec	643	30	61.2	772	6	ABR88777	AbR88777 Human sec
571	30	61.2	772	6	ABR97769	AbR97769 Human sec	644	30	61.2	772	6	ABM13429	Abm13429 Human sec
572	30	61.2	772	6	ABR87557	AbR87557 Human sec	645	30	61.2	772	6	ABM20813	Abm20813 Human sec
573	30	61.2	772	6	ABM77598	Abm77598 Human sec	646	30	61.2	772	6	ABO41944	AbO41944 Human sec
574	30	61.2	772	6	ABM27828	Abm27828 Human sec	647	30	61.2	772	6	ABO42554	AbO42554 Human sec
575	30	61.2	772	6	ABM06109	Abm06109 Human sec	648	30	61.2	772	6	ABM10074	Abm10074 Human sec
576	30	61.2	772	6	ABM03615	Abm03615 Human sec	649	30	61.2	772	6	ABO38589	AbO38589 Human sec
577	30	61.2	772	6	ABM35066	Abm35066 Human sec	650	30	61.2	772	6	ABM22643	Abm22643 Human sec
578	30	61.2	772	6	ABM26303	Abm26303 Human sec	651	30	61.2	772	6	ABM32829	Abm32829 Human sec
579	30	61.2	772	6	ABM26303	Abm26303 Human sec	652	30	61.2	772	6	ABM22643	Abm22643 Human sec
580	30	61.2	772	6	ABO48085	AbO48085 Human sec	653	30	61.2	772	6	ABM74854	Abm74854 Human sec
581	30	61.2	772	6	ABR92827	AbR92827 Human sec	654	30	61.2	772	6	ADA79650	Ada79650 Human sec
582	30	61.2	772	6	ABM11599	Abm11599 Human sec	655	30	61.2	772	6	ADA24803	Ada24803 Novel hum
583	30	61.2	772	6	ABM15996	Abm15996 Human sec	656	30	61.2	772	6	ABR96244	AbR96244 Human sec
584	30	61.2	772	6	ABM15996	Abm15996 Human sec	657	30	61.2	772	6	ABR86337	AbR86337 Human sec
585	30	61.2	772	6	ABO27557	AbO27557 Human sec	658	30	61.2	772	6	ABR86642	AbR86642 Human sec
586	30	61.2	772	6	ABO27557	AbO27557 Human sec	659	30	61.2	772	6	ABM16606	Abm16606 Human sec
587	30	61.2	772	6	ABM29048	Abm29048 Human sec	660	30	61.2	772	6	ABM29658	Abm29658 Human sec
588	30	61.2	772	6	ABM07024	Abm07024 Human sec	661	30	61.2	772	6	ABM29082	Abm29082 Human sec
589	30	61.2	772	6	ABM21118	Abm21118 Human sec	662	30	61.2	772	6	ABM23863	Abm23863 Human sec
590	30	61.2	772	6	ABM09464	Abm09464 Human sec	663	30	61.2	772	6	ADA47293	Ada47293 Human sec
591	30	61.2	772	6	ABO41334	AbO41334 Human sec	664	30	61.2	772	6	ABM22033	Abm22033 Human sec
592	30	61.2	772	6	ABO36149	AbO36149 Human PRO	665	30	61.2	772	6	ABO37674	AbO37674 Human sec
593	30	61.2	772	6	ABO43678	AbO43678 Human sec	666	30	61.2	772	6	ABM28438	Abm28438 Human sec
594	30	61.2	772	6	ABM76378	Abm76378 Human sec	667	30	61.2	772	6	ABM28743	Abm28743 Human sec
595	30	61.2	772	6	ABM76074	Abm76074 Human sec	668	30	61.2	772	6	ABM66387	Abm66387 Human sec
596	30	61.2	772	6	ABM25693	Abm25693 Human sec	669	30	61.2	772	6	ABM75769	Abm75769 Human sec
597	30	61.2	772	6	ABM25998	Abm25998 Human sec	670	30	61.2	772	6	ABM75769	Abm75769 Human sec
598	30	61.2	772	6	ABO03351	AbO03351 Human sec	671	30	61.2	772	6	ABM34049	Abm34049 Human sec
599	30	61.2	772	6	ABO02436	AbO02436 Human sec	672	30	61.2	772	6	ABM34354	Abm34354 Human sec
600	30	61.2	772	6	ABR90607	AbR90607 Human sec	673	30	61.2	772	6	ABO19682	AbO19682 Novel hum
601	30	61.2	772	6	ABR73675	AbR73675 Human sec	674	30	61.2	772	6	ABO20285	AbO20285 Human sec
602	30	61.2	772	6	ABO16927	AbO16927 Human sec	675	30	61.2	772	6	ABO21200	AbO21200 Human sec
603	30	61.2	772	6	ABR94352	AbR94352 Human sec	676	30	61.2	772	6	ABO22115	AbO22115 Human sec
604	30	61.2	772	6	ABR75859	AbR75859 Human sec	677	30	61.2	772	6	ADA12464	Ada12464 Human sec
605	30	61.2	772	6	ABR71235	AbR71235 Human sec	678	30	61.2	772	6	ABR96549	AbR96549 Human sec
606	30	61.2	772	6	ABR93132	AbR93132 Human sec	679	30	61.2	772	6	ABR85727	AbR85727 Human sec
607	30	61.2	772	6	ABR93437	AbR93437 Human sec	680	30	61.2	772	6	ABR99709	AbR99709 Human sec
608	30	61.2	772	6	ABR87862	AbR87862 Human sec	681	30	61.2	772	6		

682	30	61.2	772	6	ABM00565	Abm00565 Human sec	755	30	61.2	772	6	ABM15564	Abm15564 Human sec
683	30	61.2	772	6	ABM00260	Abm00260 Human sec	756	30	61.2	772	6	ABM08549	Abm08549 Human sec
684	30	61.2	772	6	ABO29692	AbO29692 Human sec	757	30	61.2	772	6	ABO42249	ABO42249 Human sec
685	30	61.2	772	6	ABM23558	Abm23558 Human sec	758	30	61.2	772	6	ABO37979	ABO37979 Human sec
686	30	61.2	772	6	ABM29353	Abm29353 Human sec	759	30	61.2	772	6	ABO45889	ABO45889 Human sec
687	30	61.2	772	6	ABO38284	ABO38284 Human sec	760	30	61.2	772	6	ABM66692	ABM66692 Human sec
688	30	61.2	772	6	ABO45584	ABO45584 Human sec	761	30	61.2	772	6	ABD20218	ABD20218 Human sec
689	30	61.2	772	6	ABM20508	Abm20508 Human sec	762	30	61.2	772	6	ABM19593	ABM19593 Human sec
690	30	61.2	772	6	ADA81377	Ada81377 Human sec	763	30	61.2	772	6	ABO49305	ABO49305 Human sec
691	30	61.2	772	6	ABO16622	ABO16622 Human sec	764	30	61.2	772	6	ABO49610	ABO49610 Human sec
692	30	61.2	772	6	ABO18248	ABO18248 Human sec	765	30	61.2	772	6	ADA78470	ADA78470 Human sec
693	30	61.2	772	6	ABO22675	ABO22675 Human sec	766	30	61.2	772	6	ABO19573	ABO19573 Novel hum
694	30	61.2	772	6	ABO22980	ABO22980 Human sec	767	30	61.2	772	6	ABR88167	ABR88167 Human sec
695	30	61.2	772	6	ABR93522	ABR93522 Human sec	768	30	61.2	772	6	ABM26913	ABM26913 Human sec
696	30	61.2	772	6	ABR81479	ABR81479 Human sec	769	30	61.2	772	6	ABM03310	ABM03310 Human sec
697	30	61.2	772	6	ABM77903	ABM77903 Human sec	770	30	61.2	772	6	ABO39809	ABO39809 Human sec
698	30	61.2	772	6	ABR89692	ABR89692 Human sec	771	30	61.2	772	7	ABO49915	ABO49915 Human sec
699	30	61.2	772	6	ABM26608	Abm26608 Human sec	772	30	61.2	772	7	ABO50830	ABO50830 Human sec
700	30	61.2	772	6	ABM13734	ABM13734 Human sec	773	30	61.2	772	7	ABO05286	ABO05286 Human sec
701	30	61.2	772	6	ABO28472	ABO28472 Human sec	774	30	61.2	772	7	ABR74590	ABR74590 Human sec
702	30	61.2	772	6	ABO30302	ABO30302 Human sec	775	30	61.2	772	7	ABR77069	ABR77069 Human sec
703	30	61.2	772	6	ABM07329	ABm07329 Human sec	776	30	61.2	772	7	ABM17826	ABM17826 Human sec
704	30	61.2	772	6	ABM03920	ABm03920 Human sec	777	30	61.2	772	7	ABR95877	ABR95877 Human sec
705	30	61.2	772	6	ABO37064	ABO37064 Human sec	778	30	61.2	772	7	ABO21810	ABO21810 Human sec
706	30	61.2	772	6	ABO41639	ABO41639 Human sec	779	30	61.2	772	7	ABO19980	ABO19980 Human sec
707	30	61.2	772	6	ABO35234	ABO35234 Human sec	780	30	61.2	772	7	ABO24283	ABO24283 Human sec
708	30	61.2	772	6	ABM25083	ABm25083 Human sec	781	30	61.2	772	7	ABR86032	ABR86032 Human sec
709	30	61.2	772	6	ABO47475	ABO47475 Human sec	782	30	61.2	772	7	ABM10684	ABM10684 Human sec
710	30	61.2	772	6	ABO47780	ABO47780 Human sec	783	30	61.2	772	7	ABM76683	ABM76683 Human sec
711	30	61.2	772	6	ABO48390	ABO48390 Human sec	784	30	61.2	772	7	ABR89387	ABR89387 Human sec
712	30	61.2	772	6	ABO51440	ABO51440 Human sec	785	30	61.2	772	7	ABM12514	ABM12514 Human sec
713	30	61.2	772	6	ABO51745	ABO51745 Human sec	786	30	61.2	772	7	ABM05804	ABM05804 Human sec
714	30	61.2	772	6	ABO50525	ABO50525 Human sec	787	30	61.2	772	7	ABO34929	ABO34929 Human sec
715	30	61.2	772	6	ABR79649	ABR79649 Human sec	788	30	61.2	772	7	ABM03005	ABM03005 Human sec
716	30	61.2	772	6	ABM16911	ABm16911 Human sec	789	30	61.2	772	7	ABM18983	ABM18983 Human sec
717	30	61.2	772	6	ABO17943	ABO17943 Human sec	790	30	61.2	772	7	ABM19288	ABM19288 Human sec
718	30	61.2	772	6	ABO20895	ABO20895 Human sec	791	30	61.2	772	7	ABO46499	ABO46499 Human sec
719	30	61.2	772	6	ABR96854	ABR96854 Human sec	792	30	61.2	772	7	ABO49000	ABO49000 Human sec
720	30	61.2	772	6	ABM12209	ABm12209 Human sec	793	30	61.2	772	7	ABR69043	ABR69043 Human sec
721	30	61.2	772	6	ABM16301	ABm16301 Human sec	794	30	61.2	772	7	ABR89082	ABR89082 Human sec
722	30	61.2	772	6	ABM24168	ABm24168 Human sec	795	30	61.2	772	7	ABR72455	ABR72455 Human sec
723	30	61.2	772	6	ABM14649	ABM14649 Human sec	796	30	61.2	772	7	ABR74285	ABR74285 Human sec
724	30	61.2	772	6	ABM04530	ABm04530 Human sec	797	30	61.2	772	7	ABO18553	ABO18553 Human sec
725	30	61.2	772	6	ABM06719	ABm06719 Human sec	798	30	61.2	772	7	ABR80259	ABR80259 Human sec
726	30	61.2	772	6	ABM09159	ABm09159 Human sec	799	30	61.2	772	7	ABM01480	ABM01480 Human sec
727	30	61.2	772	6	ABO39199	ABO39199 Human sec	800	30	61.2	772	7	ABM02090	ABM02090 Human sec
728	30	61.2	772	6	ABM75464	ABm75464 Human sec	801	30	61.2	772	7	ABR87252	ABR87252 Human sec
729	30	61.2	772	6	ABM25388	ABm25388 Human sec	802	30	61.2	772	7	ABM12819	ABM12819 Human sec
730	30	61.2	772	6	ABM19898	ABm19898 Human sec	803	30	61.2	772	7	ABM30573	ABM30573 Human sec
731	30	61.2	772	6	ABO48804	ABO48804 Human sec	804	30	61.2	772	7	ABM24473	ABM24473 Human sec
732	30	61.2	772	6	ABO47109	ABO47109 Human sec	805	30	61.2	772	7	ABO29387	ABO29387 Human sec
733	30	61.2	772	6	ADA83175	Ada83175 Human sec	806	30	61.2	772	7	ABO31217	ABO31217 Human sec
734	30	61.2	772	6	ABR71540	ABR71540 Human sec	807	30	61.2	772	7	ABM14344	ABM14344 Human sec
735	30	61.2	772	6	ABR72150	ABR72150 Human sec	808	30	61.2	772	7	ABM09769	ABM09769 Human sec
736	30	61.2	772	6	ABR98489	ABR98489 Human sec	809	30	61.2	772	7	ABO38894	ABO38894 Human sec
737	30	61.2	772	6	ABO06859	ABO06859 Human sec	810	30	61.2	772	7	ABM34659	ABM34659 Human sec
738	30	61.2	772	6	ABR84812	ABR84812 Human sec	811	30	61.2	772	7	ABO51135	ABO51135 Human sec
739	30	61.2	772	6	ABR73370	ABR73370 Human sec	812	30	61.2	772	7	ABO03961	ABO03961 Human sec
740	30	61.2	772	6	ABR76464	ABR76464 Human sec	813	30	61.2	772	7	ABO10431	ABO10431 Human sec
741	30	61.2	772	6	ABR73065	ABR73065 Human sec	814	30	61.2	772	7	ABR77674	ABR77674 Human sec
742	30	61.2	772	6	ABM18131	ABm18131 Human sec	815	30	61.2	772	7	ABR78884	ABR78884 Human sec
743	30	61.2	772	6	ABO20590	ABO20590 Human sec	816	30	61.2	772	7	ABO23978	ABO23978 Human sec
744	30	61.2	772	6	ABO25333	ABO25333 Human sec	817	30	61.2	772	7	ABR93742	ABR93742 Human sec
745	30	61.2	772	6	ABO25638	ABO25638 Human sec	818	30	61.2	772	7	ABM01785	ABM01785 Human sec
746	30	61.2	772	6	ABR94047	ABR94047 Human sec	819	30	61.2	772	7	ABM78208	ABM78208 Human sec
747	30	61.2	772	6	ABR79954	ABR79954 Human sec	820	30	61.2	772	7	ABR89997	ABR89997 Human sec
748	30	61.2	772	6	ABM11294	ABm11294 Human sec	821	30	61.2	772	7	ABM27523	ABM27523 Human sec
749	30	61.2	772	6	ABO32901	ABO32901 Human sec	822	30	61.2	772	7	ABM13124	ABM13124 Human sec
750	30	61.2	772	6	ABO30607	ABO30607 Human sec	823	30	61.2	772	7	ABO31827	ABO31827 Human sec
751	30	61.2	772	6	ABO30912	ABO30912 Human sec	824	30	61.2	772	7	ABM14039	ABM14039 Human sec
752	30	61.2	772	6	ABM27218	ABM27218 Human sec	825	30	61.2	772	7	ABM08244	ABM08244 Human sec
753	30	61.2	772	6	ABM29963	ABM29963 Human sec	826	30	61.2	772	7	ABO40114	ABO40114 Human sec
754	30	61.2	772	6	ABM05499	ABm05499 Human sec	827	30	61.2	772	7	ABM74549	ABM74549 Human sec

828	30	61.2	772	7	ABM33744	Abm33744 Human sec	901	30	61.2	772	8	ADP40655	Adf40655 Human sec
829	30	61.2	772	7	ABM20203	Abm20203 Human sec	902	30	61.2	772	8	ADP23599	Adf23599 Human sec
830	30	61.2	772	7	ABO48695	AbO48695 Human sec	903	30	61.2	772	8	ADP33582	Adf33582 Human sec
831	30	61.2	772	7	ABR72760	ABr72760 Human sec	904	30	61.2	772	8	ADP27049	Adf27049 Human sec
832	30	61.2	772	7	ABO15402	AbO15402 Human sec	905	30	61.2	772	8	ADP27685	Adf27685 Human sec
833	30	61.2	772	7	ABR85117	ABr85117 Human sec	906	30	61.2	772	8	ADP41279	Adf41279 Human sec
834	30	61.2	772	7	ABO15097	AbO15097 Human sec	907	30	61.2	772	8	ADP32958	Adf32958 Human sec
835	30	61.2	772	7	ABO17232	AbO17232 Human sec	908	30	61.2	772	8	ADP25324	Adf25324 Human sec
836	30	61.2	772	7	ABO19867	AbO19867 Human sec	909	30	61.2	772	8	ADP26425	Adf26425 Human sec
837	30	61.2	772	7	ABO19867	AbO19867 Human sec	910	30	61.2	772	8	ADP34214	Adf34214 Human sec
838	30	61.2	772	7	ABR85422	ABr85422 Human sec	911	30	61.2	772	8	ADP46451	Adf46451 Human sec
839	30	61.2	772	7	ABR6988	ABr6988 Human sec	912	30	61.2	772	8	ADP96068	Adf96068 Novel hum
840	30	61.2	772	7	ABO28167	AbO28167 Human sec	913	30	61.2	772	8	ADG04339	Adg04339 Novel hum
841	30	61.2	772	7	ABM22948	ABm22948 Human sec	914	30	61.2	772	8	ADG00499	Adg00499 Novel hum
842	30	61.2	772	7	ABM30268	ABm30268 Human sec	915	30	61.2	772	8	ADG82755	Adg82755 Human PRO
843	30	61.2	772	7	ABM21728	ABm21728 Human sec	916	30	61.2	772	8	ADH26036	Adh26036 Novel hum
844	30	61.2	772	7	ABM21423	ABm21423 Human sec	917	30	61.2	772	8	ADH50437	Adh50437 Human sec
845	30	61.2	772	7	ABM14954	ABm14954 Human sec	918	30	61.2	772	8	ADG63651	Adg63651 Human sec
846	30	61.2	772	7	ABO41029	AbO41029 Human sec	919	30	61.2	772	8	ADG49813	Adg49813 Human sec
847	30	61.2	772	7	ABO36759	AbO36759 Human sec	920	30	61.2	772	8	ADG51685	Adg51685 Human sec
848	30	61.2	772	7	ABO37369	AbO37369 Human sec	921	30	61.2	772	8	ADH33005	Adh33005 Human PRO
849	30	61.2	772	7	ABM75159	ABm75159 Human sec	922	30	61.2	772	8	ADG49189	Adg49189 Human sec
850	30	61.2	772	7	ABM33439	ABm33439 Human sec	923	30	61.2	772	8	ADG48565	Adg48565 Human sec
851	30	61.2	772	7	ABO46194	AbO46194 Human PRO	924	30	61.2	772	8	ADG51061	Adg51061 Human sec
852	30	61.2	772	7	ADA82541	ADa82541 Human sec	925	30	61.2	772	8	ADG59005	Adg59005 Human sec
853	30	61.2	772	7	ABM31793	ABm31793 Human sec	926	30	61.2	772	8	ADG62461	Adg62461 Human sec
854	30	61.2	772	7	ABM31183	ABm31183 Human sec	927	30	61.2	772	8	ADH25486	Adh25486 Human neu
855	30	61.2	772	7	ABP73770	ABp73770 Human PRO	928	30	61.2	772	8	ADJ54744	Adj54744 Human PRO
856	30	61.2	772	7	ABP85849	ABp85849 Human sec	929	30	61.2	772	8	ADJ64515	Adj64515 Human PRO
857	30	61.2	772	7	ABM32098	ABm32098 Human sec	930	30	61.2	772	8	ADM17263	Adm17263 Human sec
858	30	61.2	772	7	ABM32403	ABm32403 Human sec	931	30	61.2	772	8	ADM36458	Adm36458 Novel hum
859	30	61.2	772	7	ABM31488	ABm31488 Human sec	932	30	61.2	772	8	ADM40263	Adm40263 Novel hum
860	30	61.2	772	7	ABM30878	ABm30878 Human sec	933	30	61.2	772	8	ADM40263	Adm40263 Novel hum
861	30	61.2	772	7	ABP76486	ABp76486 Human PRO	934	30	61.2	772	8	ADL07097	Adl07097 Human sec
862	30	61.2	772	7	ADC43912	ADc43912 Human sec	935	30	61.2	772	8	ADN37871	Adn37871 Novel hum
863	30	61.2	772	7	ADC61672	ADc61672 Human sec	936	30	61.2	772	8	ADP55976	Adp55976 Human PRO
864	30	61.2	772	7	ADC66336	ADc66336 Human sec	937	30	61.2	772	8	ADU50219	Adu50219 PRO941, S
865	30	61.2	772	7	ADC66736	ADc66736 Human sec	938	30	61.2	772	9	ADW49498	Adw49498 Human sec
866	30	61.2	772	7	ADC68860	ADc68860 Human sec	939	30	61.2	772	9	ADZ52159	Adz52159 Human sec
867	30	61.2	772	7	ADC68920	ADc68920 Human sec	940	30	61.2	774	8	ADZ828353	Adz828353 Bacterial
868	30	61.2	772	7	ADC67985	ADc67985 Human sec	941	30	61.2	774	8	ADZ44794	Adz44794 Bacterial
869	30	61.2	772	7	ADC41305	ADc41305 Human sec	942	30	61.2	789	6	ABU25481	Abu25481 Protein e
870	30	61.2	772	7	ADC67360	ADc67360 Human sec	943	30	61.2	791	6	ABU48892	Abu48892 Protein e
871	30	61.2	772	7	ADC62296	ADc62296 Human sec	944	30	61.2	798	6	ABU38404	Abu38404 Protein e
872	30	61.2	772	7	ADC41929	ADc41929 Human sec	945	30	61.2	798	6	ABU40004	Abu40004 Protein e
873	30	61.2	772	7	ADP05579	ADp05579 Human sec	946	30	61.2	798	8	ADZ24706	Adz24706 Bacterial
874	30	61.2	772	7	ADP49298	ADp49298 Human sec	947	30	61.2	799	8	ADZ27901	Adz27901 Bacterial
875	30	61.2	772	7	ADP35352	ADp35352 Human sec	948	30	61.2	805	6	ABU22655	Abu22655 Protein e
876	30	61.2	772	7	ADP16466	ADp16466 Human sec	949	30	61.2	807	6	ABU21205	Abu21205 Protein e
877	30	61.2	772	7	ADD73081	ADd73081 Human sec	950	30	61.2	807	6	ABU19894	Abu19894 Protein e
878	30	61.2	772	7	ADD72439	ADd72439 Human sec	951	30	61.2	808	8	ADN74391	Adn74391 Thale cre
879	30	61.2	772	7	ADP17090	ADp17090 Human sec	952	30	61.2	811	3	ABG35867	Abg35867 Bacterial
880	30	61.2	772	7	ADP47104	ADp47104 Human sec	953	30	61.2	811	3	ABG35867	Abg35867 Bacterial
881	30	61.2	772	7	ADG02574	ADg02574 Novel hum	954	30	61.2	820	6	ABP79190	Abp79190 N. gonorr
882	30	61.2	772	7	ADG01281	ADg01281 Novel hum	955	30	61.2	820	6	ABU37137	Abu37137 Protein e
883	30	61.2	772	7	ADG95456	ADg95456 Novel hum	956	30	61.2	820	6	ABU37137	Abu37137 Protein e
884	30	61.2	772	7	ADG12271	ADg12271 Novel hum	957	30	61.2	820	6	ABU37980	Abu37980 Pseudomon
885	30	61.2	772	7	ADH08931	ADh08931 Human PRO	958	30	61.2	820	7	ABO81458	ABo81458 Pseudomon
886	30	61.2	772	7	ADG52861	ADg52861 Human sec	959	30	61.2	820	8	ADP08282	Adp08282 Neisseria
887	30	61.2	772	7	ADG60181	ADg60181 Human sec	960	30	61.2	824	7	ADB64361	ADb64361 Human pro
888	30	61.2	772	7	ADG63802	ADg63802 Human sec	961	30	61.2	826	6	ABU28255	Abu28255 Protein e
889	30	61.2	772	7	ADJ32712	ADj32712 Novel hum	962	30	61.2	828	5	ABP69286	Abp69286 Human pol
890	30	61.2	772	7	ADM30246	ADm30246 Novel hum	963	30	61.2	828	6	ABP58225	Abp58225 Human cel
891	30	61.2	772	7	ADT60941	ADt60941 Human sec	964	30	61.2	842	8	ADQ65365	Adq65365 Novel hum
892	30	61.2	772	8	ADP48598	ADp48598 Human sec	965	30	61.2	848	5	ABG79689	Abg79689 Tumour in
893	30	61.2	772	8	ADP74243	ADp74243 Human sec	966	30	61.2	855	7	ABM90485	ABm90485 Rice abio
894	30	61.2	772	8	ADE74216	ADe74216 Human sec	967	30	61.2	859	4	ABG21946	ABg21946 Novel hum
895	30	61.2	772	8	ADP74855	ADp74855 Human sec	968	30	61.2	881	6	ABU48605	Abu48605 Protein e
896	30	61.2	772	8	ADP89699	ADp89699 Human sec	969	30	61.2	889	8	ADN23791	Adn23791 Bacterial
897	30	61.2	772	8	ADP61339	ADp61339 Human sec	970	30	61.2	900	8	ADP26907	Adp26907 Human N-C
898	30	61.2	772	8	ADP40031	ADp40031 Human sec	971	30	61.2	905	5	ABG93053	ABg93053 S. cerevi
899	30	61.2	772	8	ADP45827	ADp45827 Human sec	972	30	61.2	905	6	ABR53130	ABr53130 Protein s
900	30	61.2	772	8	ADP24223	ADp24223 Human sec	973	30	61.2	905	7	ADK62564	ADk62564 Disease t

974 30 61.2 905 8 ADQ20653
975 30 61.2 906 3 AAY70741
976 30 61.2 906 4 ABG14316
977 30 61.2 906 5 ABB81474
978 30 61.2 906 5 ABB57233
979 30 61.2 906 6 ABR58643
980 30 61.2 906 6 ABR47406
981 30 61.2 906 7 ADE55478
982 30 61.2 906 7 ADE55482
983 30 61.2 906 7 ADE55488
984 30 61.2 906 7 ADE55476
985 30 61.2 906 7 ADE55486
986 30 61.2 906 7 ADE55484
987 30 61.2 906 7 ADE55490
988 30 61.2 906 7 ADE55480
989 30 61.2 906 7 ADN95503
990 30 61.2 906 8 ADR69159
991 30 61.2 906 8 ADN23873
992 30 61.2 906 9 ADY70663
993 30 61.2 906 9 ADZ26571
994 30 61.2 912 5 ABB81471
995 30 61.2 913 8 ADU73677
996 30 61.2 916 7 ADN38968
997 30 61.2 916 8 ADU73673
998 30 61.2 922 3 AAG35866
999 30 61.2 947 4 ABB59355
1000 30 61.2 976 5 ABP35659

ALIGNMENTS

RESULT 1
AAY22358
ID AAY22358 standard; peptide; 10 AA.
XX AC AAY22358;
XX DT 27-SEP-1999 (first entry)
XX DE TPO receptor binding peptide sequence, SEQ ID NO. 9.
XX KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopoenaia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX OS Synthetic.
XX PN US5932546-A.
XX PD 03-AUG-1999.
XX PF 04-OCT-1996; 96US-00726464.
XX PR 04-OCT-1996; 96US-00726464.
XX PA (GLAXO) GLAXO WELLCOME INC.
XX PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX WPI; 1999-457122/38.
XX New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopoenaia.
XX Disclosure; Col 13-14; 36pp; English.
XX This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopoenaia,

CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC accelerate thrombocyte regeneration
XX SQ Sequence 10 AA;
Query Match 100.0%; Score 49; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
Db 1 ESVREQVMKY 10
RESULT 2
AAB16982
ID AAB16982 standard; peptide; 10 AA.
XX AC AAB16982;
XX DT 31-OCT-2000 (first entry)
XX DE TPO-mimetic peptide sequence SEQ ID NO:38.
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX OS Synthetic.
XX PN WO200024782-A2.
XX PD 04-MAY-2000.
XX PF 25-OCT-1999; 99WO-US025044.
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX Claim 19; Page 208; 608pp; English.
XX The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)-a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 10 AA;
 Query Match 100.0%; Score 49; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 |||||
 Db 1 ESVREQVMKY 10

RESULT 3
 ABB72868
 ID ABB72868 standard; peptide; 10 AA.
 XX
 AC ABB72868;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE TPO mimetic peptide SEQ ID NO:38.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IGG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX WO200183525-A2.
 PN
 XX
 XX 08-NOV-2001.
 PD
 XX
 XX 02-MAY-2001; 2001WO-US014310.
 PF
 XX
 XX 03-MAY-2000; 2000US-00563286.
 PR
 XX
 XX (AMGE-) AMGEN INC.
 PA
 XX
 XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 PI
 XX
 XX WPI; 2002-130313/17.
 DR
 XX
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 XX
 XX Claim 39; Page 43; 176pp; English.

XX
 XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 10 AA;
 Query Match 100.0%; Score 49; DB 5; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 |||||
 Db 1 ESVREQVMKY 10

RESULT 4
 ADJ73019
 ID ADJ73019 standard; peptide; 10 AA.
 XX
 AC ADJ73019;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE TPO mimetic peptide sequence SeqID 473.
 XX
 KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
 KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
 KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
 KW TPO.
 XX
 XX Synthetic.
 OS
 XX WO2003084477-A2.
 PN
 XX
 XX 16-OCT-2003.
 PD
 XX
 XX 24-MAR-2003; 2003WO-US009139.
 PF
 XX
 XX 29-MAR-2002; 2002US-0368791P.
 PR
 XX
 XX (CENZ) CENTOCOR INC.
 PA
 XX
 XX Heavner GA, Knight DM, Scallan BJ, Ghrayeb J;
 PI
 XX
 XX WPI; 2003-804237/75.
 DR
 XX
 XX New CDR mimetibody comprising a portion of a heavy or light chain
 PT variable region comprising human framework or ligand binding region,
 PT useful for preparing a composition for treating e.g., immune,
 PT cardiovascular or neurologic disease.
 XX
 XX Disclosure; SEQ ID NO 473; 97pp; English.

XX
 XX This invention relates to novel mammalian CDR mimetibodies, specific
 CC portions or variants thereof. Specifically, it refers to an antibody
 CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 49; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
Db 1 ESVREQVMKY 10

RESULT 5
ADJ52654
ID ADJ52654 standard; peptide; 10 AA.

XX AC ADJ52654;

XX DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID473.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nontropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX OS Unidentified.
OS Synthetic.

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 473; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nontropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 10 AA;
Query Match 100.0%; Score 49; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 1 ESVREQVMKY 10

RESULT 6

ADJ51615

ID ADJ51615 standard; peptide; 10 AA.

XX AC ADJ51615;

XX DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID473.

XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunologic disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX OS Unidentified.

OS Synthetic.

XX PN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;

XX DR WPI; 2004-082872/08.

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.

XX Claim 14; SEQ ID NO 473; 123pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetoid, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetoid of the invention.

XX Sequence 10 AA;

Query Match 100.0%; Score 49; DB 8; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

Db 1 ESVREQVMKY 10

RESULT 7

ABU36234

ID ABU36234 standard; protein; 291 AA.

XX AC ABU36234;

XX DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #21761.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Mycoplasma pneumoniae.

XX WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815249.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA40104.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 64159; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 291 AA;

Query Match 77.6%; Score 38; DB 6; Length 291;

Best Local Similarity 60.0%; Pred. No. 55;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

Db 191 EALREQLIKY 200

RESULT 8

ADN99878

ID ADN99878 standard; protein; 216 AA.

XX AC ADN99878;

XX DT 29-JUL-2004 (first entry)

XX DE Novel human protein sequence #694.

XX anti-inflammatory; dermatological; neuroprotective; immunomodulator;
 KW antibacterial; virucide; antipsoriatic; cytostatic; gene therapy;
 KW vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
 KW early aging; hormonal imbalance; ischemic heart disease;
 KW ulcerative colitis.

XX OS Homo sapiens.

XX PN WO2004038003-A2.

XX PD 06-MAY-2004.

XX PF 24-OCT-2003; 2003WO-US033947.

XX PR 25-OCT-2002; 2002US-0421061P.

XX PR 25-OCT-2002; 2002US-0421080P.

XX PR 25-OCT-2002; 2002US-0421552P.

XX PR 25-OCT-2002; 2002US-0421614P.

XX PR 30-OCT-2002; 2002US-0422177P.

XX PR 30-OCT-2002; 2002US-0422178P.

XX PR 15-NOV-2002; 2002US-0426355P.

XX PR 15-NOV-2002; 2002US-0426384P.

XX PR 15-NOV-2002; 2002US-0426394P.

XX PR 15-NOV-2002; 2002US-0428430P.

XX PR 15-NOV-2002; 2002US-0428916P.

XX PR 27-NOV-2002; 2002US-0429224P.

XX PR 27-NOV-2002; 2002US-0429275P.

XX PR 27-NOV-2002; 2002US-0429302P.

XX PR 27-NOV-2002; 2002US-0429326P.

XX PR 27-NOV-2002; 2002US-0429651P.

XX PR 04-DEC-2002; 2002US-0430645P.

XX PR 04-DEC-2002; 2002US-0430651P.

PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440820P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
XX PI Wong JGP, Wu G, Zhang H, Zeng C;
XX WPI; 2004-365511/34.
XX DR N-PSDB; ADN99094.
XX
XX New nucleic acid molecules, useful in preparing a composition for
PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
PT ulcerative colitis.
XX
XX Claim 14; SEQ ID NO 1478; 532pp; English.
XX
XX The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder,
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC protein of the invention.
XX
XX Sequence 216 AA;
XX
XX Query Match 75.5%; Score 37; DB 8; Length 216;

Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
Db 72 EKIREQIKKY 81
RESULT 9
AAY20122
ID AAY20122 standard; protein; 384 AA.
XX AC AAY20122;
XX DT 19-JUL-1999 (first entry)
XX DE B. burgdorferi antigenic protein, f6-27.nt.
XX KW Antigenic protein; vaccine; Lyme disease; infection; detection.
XX OS Borrelia burgdorferi.
XX PN W09859071-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US012718.
XX PR 20-JUN-1997; 97US-0050359P.
XX PR 22-JUL-1997; 97US-0053344P.
XX PR 22-JUL-1997; 97US-0053377P.
XX PR 03-SEP-1997; 97US-0057483P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MED1-) MEDIMMUNE INC.
XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX WPI; 1999-189980/16.
XX DR N-PSDB; AAX61819.
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases caused
PT by Borrelia, particularly Lyme disease.
XX PS Claim 12; Page 206; 275pp; English.
XX CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus
XX SQ Sequence 384 AA;
XX
XX Query Match 75.5%; Score 37; DB 2; Length 384;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
Db 33 ESLREQVMKY 42
RESULT 10
AAY06993
ID AAY06993 standard; protein; 429 AA.
XX AC AAY06993;
XX DT 02-JUL-1999 (first entry)

XX DE Rat eukaryotic translation initiation factor (eIF-5).

XX KW Cancer associated antigen; diagnosis; research; treatment; human;

XX KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;

XX KW prostate cancer.

XX OS Rattus sp.

XX XX WO9904265-A2.

XX XX 28-JAN-1999.

XX XX 15-JUL-1998; 98WO-US014679.

XX XX 17-JUL-1997; 97US-00896164.

XX PR 10-OCT-1997; 97US-0061599P.

XX PR 10-OCT-1997; 97US-0061765P.

XX PR 10-OCT-1997; 97US-00948705.

XX PR 11-OCT-1997; 97GB-00021697.

XX PR 22-JUN-1998; 98US-00102322.

XX XX (LUDW-) LUDWIG INST CANCER RES.

XX PA Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;

XX PI O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;

XX XX WPI; 1999-132448/11.

XX DR N-PSDB; AAX40194.

XX XX New isolated cancer associated nucleic acids and polypeptides - isolated

XX PT using sera from cancer patients, used to develop products for the

XX PT diagnosis, monitoring or treatment of cancers.

XX XX Example 8; Page 771-772; 787pp; English.

XX CC The invention relates to a method for diagnosing a disorder characterised

XX CC by expression of a human cancer associated antigen precursor coded for by

XX CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a

XX CC biological sample isolated from a subject with an agent that specifically

XX CC binds to the NAM, an expression product or a fragment of an expression

XX CC product complexed with an HLA molecule; and (b) determining the

XX CC interaction between the agent and the NAM or the expression product as a

XX CC determination of the disorder. The products and methods can be used in

XX CC the diagnosis, monitoring, research, or treatment of conditions

XX CC characterised by the expression of various cancer associated antigens.

XX CC The invention provides nucleic acid sequences and encoded polypeptides

XX CC which are cancer associated antigen precursors expressed in human breast

XX CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and

XX CC lung cancer

XX XX Sequence 429 AA;

XX SQ Query Match 75.5%; Score 37; DB 2; Length 429;

XX Best Local Similarity 60.0%; Pred. No. 1.3e+02;

XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10

Db 285 EKIREQIKKY 294

RESULT 11

ADB60677

ID ADB60677 standard; protein; 429 AA.

XX AC ADE60677;

XX XX 29-JAN-2004 (first entry)

XX DT Rat Protein Q07205, SEQ ID NO 6589.

XX DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX KW

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX XX WO2003016475-A2.

XX PD 27-FEB-2003.

XX XX 14-AUG-2002; 2002WO-US025765.

XX XX 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX XX Woolf C, D'urso D, Befort K, Costigan M;

XX PI WPI; 2003-268312/26.

XX DR GENBANK; Q07205.

XX DR New composition comprising two or more isolated polypeptides, useful for

XX PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

XX CC or human polynucleotides or a polynucleotide which represents a fragment,

XX CC derivative or allelic variation of the nucleic acid sequence. Also

XX CC claimed are a vector comprising the novel polynucleotide, a host cell

XX CC comprising the vector, a method for identifying a nucleotide sequence

XX CC which is differentially regulated in an animal subjected to pain and a

XX CC kit to perform the method, an array, a method for identifying an agent

XX CC that increases or decreases the expression of the polynucleotide sequence

XX CC that is differentially expressed in neuronal tissue of a first animal

XX CC subjected to pain, a method for identifying a compound which regulates

XX CC the expression of a polynucleotide sequence which is differentially

XX CC expressed in an animal subjected to pain, a method for identifying a

XX CC compound that regulates the activity of one or more of the

XX CC polynucleotides, a method for producing a pharmaceutical composition, a

XX CC method for identifying a compound or small molecule that regulates the

XX CC activity in an animal of one or more of the polypeptides given in the

XX CC specification, a method for identifying a compound useful in treating

XX CC pain and a pharmaceutical composition comprising the one or more

XX CC polypeptides or their antibodies. The polynucleotide or the compound that

XX CC modulates its activity is useful for preparing a medicament for treating

XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of

XX CC the specification) which is differentially expressed during pain. Note:

XX CC The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic form directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences.

XX XX Sequence 429 AA;

XX SQ Query Match 75.5%; Score 37; DB 7; Length 429;

XX Best Local Similarity 60.0%; Pred. No. 1.3e+02;

XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10

Db 285 EKIREQIKKY 294

RESULT 12

ADD45400

ID ADD45400 standard; protein; 429 AA.

XX AC ADD45400;

XX XX 29-JAN-2004 (first entry)

XX DT

```
XX DE Rat Protein Q07205, SEQ ID NO 10833.
XX AC
XX DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX DE
XX OS Rattus norvegicus.
XX KW WO2003016475-A2.
XX PN
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI
XX PF Woolf C, D'urso D, Befort K, Costigan M;
XX XX WPI; 2003-268312/26.
XX DR GENBANK; Q07205.
XX DR
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX XX
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 429 AA;

Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294

RESULT 13
ADE60681
ID ADE60681 standard; protein; 429 AA.
```

```
XX AC ADE60681;
XX AC
XX DT 29-JAN-2004 (first entry)
XX DE
XX DE Rat Protein Q07205, SEQ ID NO 6593.
XX XX
XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX OS Rattus norvegicus.
XX KW WO2003016475-A2.
XX PN
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PA (FARB ) BAYER AG.
XX PI
XX PF Woolf C, D'urso D, Befort K, Costigan M;
XX XX WPI; 2003-268312/26.
XX DR GENBANK; Q07205.
XX DR
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX XX
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 429 AA;

Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
```

RESULT 14
ADE61162
ID ADE61162 standard; protein; 429 AA.
XX AC ADE61162;
XX AC
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Rat Protein Q07205, SEQ ID NO 7080.
XX XX
XX XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX KW Rattus norvegicus.
XX OS
XX XX WO2003016475-A2.
XX PN
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX XX (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX XX
XX DR WPI; 2003-268312/26.
XX DR GENBANK; Q07205.
XX XX
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX XX
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 429 AA;
Query Match 75.5%; Score 37; DB 7; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ESVREQVMKY 10
| : ||| : ||
Db 285 EKIREQIKKY 294
RESULT 15
ADE61164
ID ADE61164 standard; protein; 431 AA.
XX AC ADE61164;
XX AC ADE61164;
XX XX
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Human Protein P55010, SEQ ID NO 7082.
XX XX
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX XX
XX PN WO2003016475-A2.
XX XX
XX PD 27-FEB-2003.
XX XX
XX PF 14-AUG-2002; 2002WO-US025765.
XX XX
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX XX (GEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX XX
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX XX
XX DR WPI; 2003-268312/26.
XX DR GENBANK; P55010.
XX XX
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.
XX XX
XX CC The invention discloses a composition comprising two or more isolated rat
XX CC or human polynucleotides or a polynucleotide which represents a fragment,
XX CC derivative or allelic variation of the nucleic acid sequence. Also
XX CC claimed are a vector comprising the novel polynucleotide, a host cell
XX CC comprising the vector, a method for identifying a nucleotide sequence
XX CC which is differentially regulated in an animal subjected to pain and a
XX CC kit to perform the method, an array, a method for identifying an agent
XX CC that increases or decreases the expression of the polynucleotide sequence
XX CC that is differentially expressed in neuronal tissue of a first animal
XX CC subjected to pain, a method for identifying a compound which regulates
XX CC the expression of a polynucleotide sequence which is differentially
XX CC expressed in an animal subjected to pain, a method for identifying a
XX CC compound that regulates the activity of one or more of the
XX CC polynucleotides, a method for producing a pharmaceutical composition, a
XX CC method for identifying a compound or small molecule that regulates the
XX CC activity in an animal of one or more of the polypeptides given in the
XX CC specification, a method for identifying a compound useful in treating
XX CC pain and a pharmaceutical composition comprising the one or more
XX CC polypeptides or their antibodies. The polynucleotide or the compound that
XX CC modulates its activity is useful for preparing a medicament for treating
XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX CC therapy). The sequence presented is a human protein (shown in Table 2 of
XX CC the specification) which is differentially expressed during pain. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic form directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 16
ID ADE60679 standard; protein; 431 AA.
XX
XX ADE60679;
XX
XX 29-JAN-2004 (first entry)
XX Human Protein P55010, SEQ ID NO 6591.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

Qy 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 17
ID ADD45402 standard; protein; 431 AA.
XX
XX ADD45402;
XX
XX 29-JAN-2004 (first entry)
XX Human Protein P55010, SEQ ID NO 10835.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P55010.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:

CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 18

ADBE60683
ID ADBE60683 standard; protein; 431 AA.

XX ADE60683;

XX 29-JAN-2004 (first entry)

XX Human Protein P55010, SEQ ID NO 6595.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.
DR GENBANK; P55010.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 431 AA;

Query Match 75.5%; Score 37; DB 7; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 19

ADN04323
ID ADN04323 standard; protein; 431 AA.

XX ADN04323;

XX 01-JUL-2004 (first entry)

XX Antipsoriatic protein sequence #356.

XX antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

XX Wu TD;

XX WPI; 2004-305105/28.

XX N-PSDB; ADN04322.

XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.

XX Claim 9; SEQ ID NO 717; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.

XX Sequence 431 AA;

Query Match 75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY      1  ESVREQVMKY 10
Db      287 EKIREQIKKY 296

RESULT 20
ID  ABM81587 standard; protein; 431 AA.
XX
AC  ABM81587;
XX
DT  18-NOV-2004 (first entry)
XX
DE  Tumour-associated antigenic target (TAT) polypeptide PRO82391, SEQ:4099.
XX
KW  Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW  tumour; diagnosis; cell proliferative disorder; breast cancer;
KW  colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW  central nervous system cancer; bladder cancer; pancreatic cancer;
KW  cervical cancer; melanoma; leukaemia; hybridisation probe;
KW  chromosome identification; chromosome mapping; gene mapping;
KW  gene therapy; cytostatic.
XX
OS  Homo sapiens.
XX
PN  WO2004030615-A2.
XX
PD  15-APR-2004.
XX
PF  29-SEP-2003; 2003WO-US028547.
XX
PR  02-OCT-2002; 2002US-0414971P.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Wu TD, Zhang Z, Zhou Y;
XX
DR  WPI; 2004-347921/32.
XX
DR  N-PSDB; ACN39755.
XX
PT  New tumor-associated antigenic target polypeptides and nucleic acids,
PT  useful in preparing a medicament for treating or detecting a
PT  proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT  prostate cancer or tumor.
XX
PS  Claim 12; SEQ ID NO 4099; 7273pp; English.
XX
CC  The invention relates to human tumour-associated antigenic target (TAT)
CC  polypeptides, and their related nucleic acids. The TAT polypeptides are
CC  overexpressed in cancer tissues compared to normal tissues, and may thus
CC  serve as effective targets for the diagnosis and treatment of cancer in
CC  mammals. The invention also relates to nucleic acid and polypeptide
CC  sequences at least 80% identical to the TAT nucleic acids and
CC  polypeptides; expression vectors and host cells comprising a TAT nucleic
CC  acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC  molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC  TAT polypeptide; and methods and compositions for the treatment or
CC  diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC  antibodies, antagonists, binding molecules and compositions are useful
CC  for diagnosing or treating a cell proliferative disorder associated with
CC  increased TAT expression, particularly cancers such as breast cancer,
CC  colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC  cancer, pancreatic cancer, cervical cancer, cancers of the central
CC  nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC  used as hybridisation probes, in chromosome and gene mapping, in
CC  chromosome identification and in gene therapy. The present sequence
CC  represents a TAT polypeptide of the invention
XX
SQ  Sequence 431 AA;

Query Match      75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1  ESVREQVMKY 10
Db      287 EKIREQIKKY 296

RESULT 21
ID  ADP54723 standard; protein; 431 AA.
XX
AC  ADP54723;
XX
DT  18-NOV-2004 (first entry)
XX
DE  Human PRO protein sequence SEQ ID NO:699.
XX
KW  human; PRO; immune related disease; inflammatory immune response;
KW  immune response stimulation; anti-allergic; antianaemic; antiarthritic;
KW  antiasthmatic; antidiabetic; antiinflammatory; antiporiatic;
KW  antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW  haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW  nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW  virucide; gene therapy.
XX
OS  Homo sapiens.
XX
PN  WO2004039956-A2.
XX
PD  13-MAY-2004.
XX
PF  28-OCT-2003; 2003WO-US034381.
XX
PR  29-OCT-2002; 2002US-0422472P.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
XX
PI  Wood WI, Wu TD;
XX
DR  WPI; 2004-376182/35.
XX
DR  N-PSDB; ADP54722.
XX
PT  New PRO polynucleotides and polypeptides, useful in diagnosing
PT  and treating an immune related disease, e.g. systemic lupus
PT  erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT  stimulating an immune response.
XX
PS  Claim 1; SEQ ID NO 699; 3009pp; English.
XX
CC  The present invention describes an isolated PRO nucleic acid (1). Also
CC  described: (1) a vector comprising (1); (2) a host cell comprising the
CC  vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC  isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC  polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC  antibody which specifically binds to a polypeptide of (4); (7) a
CC  composition of matter comprising a polypeptide of (4); an agonist or
CC  antagonist of the polypeptide or an antibody that binds to the
CC  polypeptide in combination with a carrier; (8) an article of manufacture
CC  comprising a container, a label on the container and a composition of
CC  matter of (7); (9) a method of treating an immune related disease in a
CC  mammal; (10) a method for determining the presence of a PRO polypeptide
CC  in a sample suspected of having the polypeptide; (11) a method of
CC  diagnosing an immune related disease or an inflammatory immune response
CC  in mammal; (12) a method of identifying a compound that inhibits or
CC  mimics the activity of or expression of a gene encoding a PRO polypeptide
CC  ; and (13) a method of stimulating the immune response in a mammal. The
CC  PRO sequences have anti-allergic, antianaemic, antiarthritic,
CC  antiasthmatic, antidiabetic, antiinflammatory, antiporiatic,
CC  antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC  haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC  nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC  virucide activities, and can be used in gene therapy. The nucleic acid
CC  (1) and the encoded polypeptides, compositions, kits and methods are

```

CC useful in diagnosing and treating an immune related disease and in
CC stimulating an immune response. The present sequence represents a human
CC PRO protein from the present invention.

XX SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 8; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 22

ID ADX05561
ADX05561 standard; protein; 431 AA.

XX AC ADX05561;

XX DT 21-APR-2005 (first entry)

XX DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 126.

XX KW cytotstatic; cyclin-dependent kinase; cdk; biomarker.

XX OS Homo sapiens.

XX PN WO2005012875-A2.

XX PD 10-FEB-2005.

XX PF 29-JUL-2004; 2004WO-US024424.

XX PR 29-JUL-2003; 2003US-0490890P.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;

XX WPI: 2005-163068/17.

XX DR N-PSDB; ADX05560.

XX PS Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.

XX Claim 5; SEQ ID NO 126; 141pp; English.

XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC sequence represents a biomarker used in the method of the invention.

XX SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 9; Length 431;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 23

ID ADY15998
ADY15998 standard; protein; 431 AA.

XX AC ADY15998;

XX DT 05-MAY-2005 (first entry)

XX DE PRO polypeptide SEQ ID NO 1804.

XX KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.

XX OS Homo sapiens.

XX PN WO2005016962-A2.

XX PD 24-FEB-2005.

XX PF 11-AUG-2004; 2004WO-US026249.

XX PR 11-AUG-2003; 2003US-0493546P.

XX PA (GETH) GENENTECH INC.

XX PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

XX WPI: 2005-182330/19.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

XX Claim 8; SEQ ID NO 1804; 158pp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody, an immune
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.

XX SQ Sequence 431 AA;

Query Match 75.5%; Score 37; DB 9; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ESVREQVMKY 10
| : ||| : ||
Db 287 EKIREQIKKY 296

RESULT 24

ID AAB43508
AAB43508 standard; protein; 433 AA.

XX AC AAB43508;

XX DT 08-FEB-2001 (first entry)

XX DE Human cancer associated protein sequence SEQ ID NO:953.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnarary; immunomodulator;
KW antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
XX
XX WO200055350-A1.
XX
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005882.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-587533/55.
XX
XX N-PSDB; AAC77717.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX
XX
XX Claim 11; Page 1519-1521; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnarary; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
XX nootropic; vasotropic; antipsoriatic and antiangiogenic. The
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancers, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX AAC78457 and AAB44240 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 433 AA;
XX
XX
XX Query Match 75.5%; Score 37; DB 3; Length 433;
XX Best Local Similarity 60.0%; Pred. No. 1.3e+02;
XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX Qy 1 ESVREQVWKY 10
XX | : | | | : | |
XX 289 EKIREQINKY 298
XX
XX
XX
XX RESULT 25
XX ADT56493
XX ID ADT56493 standard; protein; 514 AA.
XX
XX AC ADT56493;
XX
XX

DT 13-JAN-2005 (first entry)
XX
XX Plant polypeptide, SEQ ID 6570.
XX
XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
XX Viridiplantae.
OS
XX
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
XX
XX 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
XX genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.
XX
XX Claim 2; SEQ ID NO 6570; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
XX polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
XX 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
XX (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
XX Arabidopsis, wheat and rape but the specification does not indicate which
XX sequences is derived from which organism. Also included is a method of
XX producing a plant having an improved property, comprising transforming a
XX plant with a recombinant DNA construct comprising a promoter region
XX functional in a plant cell operably joined to a polynucleotide encoding a
XX polypeptide associated with the property, and growing the transformed
XX plant. The property is selected from improving plant cold tolerance, for
XX manipulating growth rate in plant cells by modification of the cell cycle
XX pathway, for improving plant drought tolerance, for providing increased
XX resistance to plant disease, for galactomannan production, for production
XX of plant growth regulators, for improving plant heat tolerance, for
XX improving plant tolerance to herbicides, for increasing the rate of
XX homologous recombination in plants, for lignin production, for improving
XX plant tolerance to extreme osmotic conditions, for improving plant
XX tolerance to pathogens or pests, for yield improvement by modification of
XX photosynthesis, for modifying seed oil yield and/or content, for
XX modifying seed protein yield and/or content, for yield improvement by
XX modification of carbohydrate, nitrogen or phosphorus use and/or uptake
XX and for yield improvement by providing improved plant growth and
XX development under at least one stress condition. The polynucleotide may
XX also encode a plant transcription factor. The methods and compositions of
XX the present invention are useful in the field of biochemistry and
XX genetics, in particular for producing transgenic plants with improved
XX biological characteristics such as increased yield, improved nitrogen
XX flow, increasing plant tolerance to cold or heat, improving plant
XX tolerance to extreme osmotic and drought conditions, and improving plant
XX tolerance to plant pests or pathogens. They can also be used in physical
XX arrays of molecules, plant breeding markers, computer-based storage and
XX analysis systems. The present sequence is one of the 5544 plant protein
XX sequences of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docID=20040216190.
XX
XX Sequence 514 AA;
XX
XX

Query Match 75.5%; Score 37; DB 8; Length 514;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 436 EEIREQALKY 445

RESULT 26
AAU30330
ID AAU30330 standard; protein; 590 AA.
XX
AC AAU30330;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #821.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX
PS Claim 20; Page 278; 765pp; English.

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid sequences of novel human secreted proteins of the invention
XX
SQ Sequence 590 AA;

Query Match 75.5%; Score 37; DB 4; Length 590;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 435 EKIREQIKKY 444

RESULT 27
AAY22378
ID AAY22378 standard; peptide; 15 AA.

XX
AC AAY22378;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 29.

XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopoenaia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.

XX
OS Synthetic.

XX
PN US5932546-A.

XX
PD 03-AUG-1999.

XX
PF 04-OCT-1996; 96US-00726464.

XX
PR 04-OCT-1996; 96US-00726464.

XX
PA (GLAX) GLAXO WELLCOME INC.

XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;

XX
PI Cwirla SE, Johnson SS;

XX
DR WPI; 1999-457122/38.

XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopoenia.

XX
PS Disclosure; Col 13-14; 36pp; English.

XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopoenaia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; and
CC in vitro expansion of megakaryocytes and committed progenitor cells; and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration

XX
SQ Sequence 15 AA;

Query Match 73.5%; Score 36; DB 2; Length 15;
Best Local Similarity 70.0%; Pred. No. 5.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
| : ||| : ||
Db 4 ESVRQVMKXF 13

RESULT 28
ADC89126
ID ADC89126 standard; protein; 150 AA.

XX
AC ADC89126;

XX
DT 01-JAN-2004 (first entry)

XX
DE Ribosomal protein similar to FCWP1 #1342.

XX
KW Antifungal protein; ribosomal protein; FCWP1; AlyAPP;
KW plant fungal infection; Alternaria; Ascochyta; Botrytis; Cercospora;
KW Colletotrichum; Diplodia; Fusarium; Gaemanomyces; Helminthosporium;

KW Macrospora; Mycosphaerella; Nectria; Peronospora; Phoma;
 KW Phymatotrichum; Phytophthora; Plasmodia; Podosphaera; Puccinia; Puthium;
 KW Pyrenophora; Pyricularia; Pythium; Rhizoctonia; Sclerotium; Sclerotinia;
 KW Septoria; Thielaviopsis; Venturia; Verticillium.

XX Unidentified.

XX US6573361-B1.

XX 03-JUN-2003.

XX 07-DEC-2000; 2000US-00732210.

XX 07-DEC-1999; 99US-0169340P.

XX 07-DEC-1999; 99US-0169513P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Bunkers GJ, Liang J, Mittanck CA, Seale JW, Wu YS;

XX WPI; 2003-754558/71.

XX Novel antifungal protein FCWP1, isolated from *Fusarium culmorum*, useful
 for controlling fungal infections in plants.

XX Example 21; SEQ ID NO 1379; 27pp; English.

XX The invention relates to an isolated antifungal ribosomal protein from
 CC *Fusarium culmorum*, FCWP1. Also included is a fusion protein between the
 CC signal peptide of the antifungal protein AliAPP from *Alvium* and FCWP1,
 CC encoded by the nucleic acid appearing as A0887758. The FCWP1 proteins are
 CC useful for controlling fungal infections in plants, such as those caused
 CC by *Alternaria* (e.g. *Alternaria brassicicola*, *Alternaria solani*),
 CC *Ascochyta* (e.g. *Ascochyta pisi*), *Botrytis* (e.g. *Botrytis cinerea*),
 CC *Cercospora* (e.g. *Cercospora kikuchii*, *Cercospora zea-maydis*),
 CC *Colletotrichum* (e.g. *Colletotrichum lindemuthianum*), *Diplodia* (e.g.
 CC *Fusarium graminearum*, *Fusarium culmorum*, *Fusarium solani*, *Fusarium*
 CC *moniliforme*, *Fusarium roseum*), *Gaeumannomyces* (e.g. *Gaeumannomyces*
 CC *graminis* f.sp. *tritici*), *Helminthosporium* (e.g. *Helminthosporium turcicum*
 CC *graminis* f.sp. *tritici*), *Helminthosporium* (e.g. *Helminthosporium maydis*), *Macrophoma*
 CC (e.g. *Macrophoma phaseolina*, *Maganaporthe grisea*), *Mycosphaerella*
 CC (e.g. *Mycosphaerella figiensis*), *Nectria* (e.g. *Nectria haematococca*),
 CC *Peronospora* (e.g. *Peronospora manshurica*, *Peronospora tabacina*), *Phoma*
 CC (e.g. *Phoma betae*), *Phymatotrichum* (e.g. *Phymatotrichum omnivorum*),
 CC *Phytophthora* (e.g. *Phytophthora cinnamomi*, *Phytophthora cactorum*,
 CC *Phytophthora phaseoli*, *Phytophthora parasitica*, *Phytophthora*
 CC *citrophthora*, *Phytophthora megasperma* f.sp. *sojae*, *Phytophthora*
 CC *infestans*), *Plasmodia* (e.g. *Plasmodia viticola*), *Podosphaera* (e.g.
 CC *Podosphaera leucotricha*), *Puccinia* (e.g. *Puccinia sorghi*, *Puccinia*
 CC *striformis*, *Puccinia graminis* f.sp. *tritici*, *Puccinia asparagi*,
 CC *Puccinia recondita*, *Puccinia arachidis*), *Puthium* (e.g. *Puthium*
 CC *aphanidermatum*), *Pyrenophora* (e.g. *Pyrenophora tritici-repentens*),
 CC *Pyricularia* (e.g. *Pyricularia oryzae*), *Pythium* (e.g. *Pythium ultimum*),
 CC *Rhizoctonia* (e.g. *Rhizoctonia solani*, *Rhizoctonia cerealis*), *Sclerotium*
 CC (e.g. *Sclerotium rolfsii*), *Sclerotinia* (e.g. *Sclerotinia sclerotiorum*),
 CC *Septoria* (e.g. *Septoria lycopersici*, *Septoria glycines*, *Septoria*
 CC *nodorum* / *Phaeosphaeria nodorum*, *Septoria tritici*), *Thielaviopsis* (e.g.
 CC *Thielaviopsis basicola*), *Uncinula* (e.g. *Uncinula necator*), *Venturia*
 CC (e.g. *Venturia inaequalis*) or *Verticillium* (e.g. *Verticillium dahliae*,
 CC *Verticillium albo-atrum*). Mutations in the proteolytic consensus
 CC sequences contained within FCWP1 provides improved stability of its
 CC antifungal activity. Also disclosed are ribosomal proteins with similar
 CC PI (>7) and molecular weight (<20kDa) to FCWP1, which may act as
 CC antifungal proteins. The present sequence represents one of the ribosomal
 CC proteins similar to FCWP1. Note: The sequence reference data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=6573361B1.

XX Sequence 150 AA;

Query Match 73.5%; Score 36; DB 7; Length 150;

Best Local Similarity 70.0%; Pred. No. 65;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ESVREQVMKY 10
 Db ||| ||:|
 30 ESVIEQIVKY 39

RESULT 29

ABG93066

ID ABG93066 standard; protein; 151 AA.

XX AC ABG93066;

XX DT 21-NOV-2002 (first entry)

XX S. cerevisiae BAX-associated protein fragment SEQ ID 90.

XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death.

XX OS Saccharomyces cerevisiae.

XX PN WO200264766-A2.

XX PD 22-AUG-2002.

XX PF 21-DEC-2001; 2001WO-EP015398.

XX PR 22-DEC-2000; 2000EP-00870318.

XX PR 04-JAN-2001; 2001EP-00870002.

XX PR 09-JAN-2001; 2001EP-00870003.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;

XX DR WPI; 2002-667002/71.

XX DR N-PSDB; ABQ76332.

XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.

XX Claim 36; Fig 1; 344pp; English.

XX This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide, immunosuppressive, virucide and
 CC vasotropic activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenous flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polypeptide associated with the Bax gene
 CC described in the disclosure of the invention

XX Sequence 151 AA;

Query Match 73.5%; Score 36; DB 5; Length 151;

Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 31 ESVIEQIVKY 40

RESULT 30
ADK64530
ID ADK64530 standard; protein; 151 AA.
XX
AC ADK64530;
XX
XX 06-MAY-2004 (first entry)
XX
XX Disease treating protein complex-derived protein #1110.
XX
XX protein complex; drug target; diagnosis.
XX
XX Unidentified.
XX
XX EP1338608-A2.
XX
XX 27-AUG-2003.
XX
XX 20-DEC-2002; 2002EP-00102902.
XX
XX 20-DEC-2001; 2001EP-00130253.
XX
XX (CELL-) CELLZOME AG.
XX
XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;
PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;
PI Michon A, Leutwein C, Rick J;
XX
XX WPI: 2003-638460/61.
DR N-PSDB; ADK64531.
XX
XX New proteins and protein complexes from eukaryotes, useful as targets in
PT drug screening, or in diagnosing or screening for the presence of a
PT disease or disorder, or a predisposition for developing a disease or
PT disorder in a subject.
XX
XX Disclosure; SEQ ID NO 2219; 13pp; English.

XX
XX The invention relates to novel protein complexes comprising a first and a
CC second protein, or its derivative, fragment, homologue or variant. The
CC proteins are selected from given protein complexes, which are not defined
CC in the specification. The variants are encoded by nucleic acids that
CC hybridize to the nucleic acids encoding the proteins under low stringency
CC conditions. The protein complexes are useful as targets for an active
CC agent of a pharmaceutical. These protein complexes are particularly
CC useful as drugs targets for the treatment or preventing of a disease or
CC disorder. The complexes and methods above are useful in diagnosing or
CC screening for the presence of a disease or disorder or a predisposition
CC for developing a disease or disorder in a subject. These are also useful
CC in screening for a drug for treatment or prevention of a disease or
CC disorder. The molecule that modulates the amount, activity or protein
CC components of the complex is useful for the manufacture of a medicament
CC for the treatment or prevention of a disease or disorder. This sequence
CC corresponds to a protein of the invention. (Note: the sequence data for
CC this patent did not form part of the printed specification but was
CC obtained from the EPO in electronic format).

XX
XX Sequence 151 AA;
SQ

Query Match 73.5%; Score 36; DB 7; Length 151;
Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 31 ESVIEQIVKY 40

RESULT 31
AAU67591
ID AAU67591 standard; protein; 339 AA.
XX
AC AAU67591;
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #28487.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
PN
XX 01-NOV-2001.
PD
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
PR
XX 02-JUN-2000; 2000US-0208841P.
PR
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
DR N-PSDB; AAS59583.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 28786; 1069pp; English.

XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 339 AA;
SQ

Query Match 73.5%; Score 36; DB 4; Length 339;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| |||:
Db 31 ESVIEQIVKY 40

```
Db      104 ESIXDQVSKY 113
RESULT 32
ABM64110
XX ABM64110 standard; protein; 339 AA.
XX
XX ABM64110;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes predicted ORF-encoded polypeptide #28786.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
XX Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barth B, Vallieue-Douglas J;
XX
XX WPI; 2003-381789/36.
XX N-PSDB; ACF64512.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 28786; 1481bp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of P. acnes polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
XX method for stimulating an immune response specific for a P. acnes
XX polypeptide and an isolated T cell population comprising T cells prepared
XX via this method; a vaccine composition (comprising P. acnes polypeptides,
XX polynucleotides, antibodies, fusion proteins, T cell populations, or
XX antigen-presenting cells that express the polypeptide); a method and kit
XX for detecting or determining the presence or absence of P. acnes in a
XX patient; and a method for inhibiting the development of P. acnes in a
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations or antigen-presenting cells that express the
XX polypeptides are useful for diagnosing, preventing or treating acne
XX vulgaris, or for stimulating an immune response specific for a P. acnes
XX protein. The polynucleotides can also be used as probes or primers for
XX nucleic acid hybridisation. The vaccine composition is useful for the
XX stimulation of an immune response against P. acnes, or for treating acne,
XX and the kit is useful for performing a diagnostic assay. The present
XX sequence represents a polypeptide predicted to be encoded by an ORF (open
XX reading frame) contained within the P. acnes polynucleotides of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 339 AA;
XX
XX Query Match 73.5%; Score 36; DB 6; Length 339;
XX Best Local Similarity 60.0%; Pred. No. 1.6e+02;
```

```
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
Db 104 ESIXDQVSKY 113
RESULT 33
ABM65793
XX ABM65793 standard; protein; 339 AA.
XX
XX ABM65793;
XX
XX 20-OCT-2003 (first entry)
XX
XX Propionibacterium acnes immunogenic polypeptide #30469.
XX
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine; immunogenic.
XX
XX Propionibacterium acnes.
XX
XX WO2003033515-A1.
XX
XX 24-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032727.
XX
XX 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barth B, Vallieue-Douglas J;
XX
XX WPI; 2003-381789/36.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.
XX
XX Claim 7; SEQ ID NO 30469; 1481bp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of P. acnes polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
XX method for stimulating an immune response specific for a P. acnes
XX polypeptide and an isolated T cell population comprising T cells prepared
XX via this method; a vaccine composition (comprising P. acnes polypeptides,
XX polynucleotides, antibodies, fusion proteins, T cell populations, or
XX antigen-presenting cells that express the polypeptide); a method and kit
XX for detecting or determining the presence or absence of P. acnes in a
XX patient; and a method for inhibiting the presence or absence of P. acnes in a
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations or antigen-presenting cells that express the
XX polypeptides are useful for diagnosing, preventing or treating acne
XX vulgaris, or for stimulating an immune response specific for a P. acnes
XX protein. The polynucleotides can also be used as probes or primers for
XX nucleic acid hybridisation. The vaccine composition is useful for the
XX stimulation of an immune response against P. acnes, or for treating acne,
XX and the kit is useful for performing a diagnostic assay. The present
XX sequence represents a specifically claimed P. acnes polypeptide which is
XX thought to contain an immunogenic region. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 339 AA;
```


Mon May 15 11:35:33 2006

Query Match 73.5%; Score 36; DB 6; Length 339;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|||||
DB 104 ESIKQVSKY 113

RESULT 34
ABG28636
ID ABG28636 standard; protein; 914 AA.
XX
AC ABG28636;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #28627.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
DR N-PSDB; AAS92823.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 58995; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

QY Sequence 914 AA;
Query Match 73.5%; Score 36; DB 4; Length 914;

Best Local Similarity 77.8%; Pred. No. 4.7e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
||:|||||
DB 334 ESVREQVMR 342

RESULT 35
ADQ97432
ID ADQ97432 standard; protein; 1250 AA.
XX
AC ADQ97432;
XX
DT 07-OCT-2004 (first entry)
XX
DE Mouse cancer associated sequence MP08-047, SEQ ID 409.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.
XX
OS Mus musculus.
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 409; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

QY Sequence 1250 AA;
Query Match 73.5%; Score 36; DB 8; Length 1250;
Best Local Similarity 80.0%; Pred. No. 6.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|||||
DB 589 ESVRAQVEY 598

RESULT 36
ADQ90750
ID ADQ90750 standard; protein; 2110 AA.
XX
AC ADQ90750;
XX
DT 21-OCT-2004 (first entry)
XX
DE Mouse acetyl-coenzyme A carboxylase - ACCL.
XX
KW mouse; carboxyltransferase domain; CT domain;
KW acetyl-coenzyme A carboxylase; acetyl-CoA carboxylase;
KW metabolic syndrome; diabetes; obesity; cardiovascular disease;

KW atherosclerosis; depression; cancer; hyperlipidaemia; dyslipidaemia;
KW hypertension; hyperuricaemia; renal dysfunction;
KW crystallisable composition; enzyme; murine; ACC1.
XX
OS Mus musculus.
XX
XX WO2004063715-A2.
XX
XX PD 29-JUL-2004.
XX
XX PF 09-JAN-2004; 2004WO-US0000585.
XX
XX PR 10-JAN-2003; 2003US-0439383P.
XX
XX PR 31-MAR-2003; 2003US-0459464P.
XX
XX PR 31-JUL-2003; 2003US-0491640P.
XX
XX PR 27-OCT-2003; 2003US-0514636P.
XX
XX PR 09-JAN-2004; 2004US-00514636.
XX
XX PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
XX PI Tong L, Zhang H, Yang Z;
XX
XX DR WPI; 2004-571486/55.
XX
XX PT New crystallizable composition comprising a carboxyltransferase domain of
PT acetyl-coenzyme A (CoA) carboxylase, for treating e.g. diabetes, obesity,
PT cardiovascular disease, atherosclerosis, or cancer.
XX
XX PS Example 4; SEQ ID NO 13; 195pp; English.
XX
XX SQ The invention comprises a crystallisable composition containing a
CC carboxyltransferase (Ct) domain of acetyl-coenzyme A (acetyl-CoA)
CC carboxylase. The composition of the invention is useful for treating:
CC metabolic syndrome, diabetes, obesity, cardiovascular disease,
CC atherosclerosis, depression, cancer, hyperlipidaemia, dyslipidaemia,
CC hypertension, hyperuricaemia and renal dysfunction. The present amino
CC acid sequence represents a mouse acetyl-CoA carboxylase of the invention.
XX
XX SQ Sequence 2110 AA;
Query Match 73.5%; Score 36; DB 8; Length 2110;
Best Local Similarity 70.0%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
DB 1254 ESVRSMVMRY 1263
RESULT 37
ID AAB86033 standard; protein; 2288 AA.
XX
XX AC AAB86033;
XX
XX DT 13-JUL-2001 (first entry)
XX
XX DE Bovine acetyl-coenzyme A carboxylase-alpha protein fragment.
XX
XX KW Acetyl-coenzyme A carboxylase-alpha; acetyl-CoA carboxylase alpha;
KW bovine; milk gland-specific promoter; Accalpha; milk production; sheep;
KW goat; fat content; genotyping; lactation.
XX
XX OS Bos taurus.
XX
XX PN DE19946173-A1.
XX
XX PD 05-APR-2001.
XX
XX PF 20-SEP-1999; 99DE-01046173.
XX
XX PR 20-SEP-1999; 99DE-01046173.
XX
XX PT

PA (BIOL-) FORSCHUNGSINSTITUT BIOLOGIE LANDWIRTSCHA.
XX
XX PI Seyfert HM;
XX
XX DR WPI; 2001-258968/27.
XX
XX DR N-PSDB; AAF88002.
XX
XX PT New nucleic acid encoding bovine acetyl coenzyme A carboxylase alpha and
PT its promoter, for milk-specific production of proteins and for regulating
PT fat content of milk.
XX
XX PS Claim 2c; Page 24-30; 44pp; German.
XX
XX CC This invention describes a novel milk gland-specific promoter of the
CC bovine acetyl coenzyme A carboxylase alpha gene (Accalpha), or its
CC fragments, which are used to control expression of foreign genes. When
CC the promoter (or the Accalpha structural gene) is replaced, at least in
CC part, by a sequence that is altered by deletion or substitution, then
CC expression of Accalpha in the milk gland is reduced and the milk produced
CC (by cattle, sheep or goats) has reduced fat content. Analysis of the
CC polymorphic 933-966 region of (Accalpha) is useful for genotyping
CC animals, producing a genotype that is directly correlated with Accalpha
CC expression during lactation and with fat content of the milk. This makes
CC it possible to generate populations of cattle that produce milk of high
CC or low fat content by classical breeding methods. Accalpha is a lactation
CC -specific, inducible promoter for expressing foreign proteins in the milk
CC and, when modified, results in milk of reduced fat content, which
CC facilitates recovery of proteins. This sequence represents a fragment
CC from bovine Accalpha which contains a fragment of exon 5A exon 6 and exon
CC 7
XX
XX SQ Sequence 2288 AA;
Query Match 73.5%; Score 36; DB 4; Length 2288;
Best Local Similarity 70.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 ESVREQVMKY 10
DB 1432 ESVRSMVMRY 1441
RESULT 38
ID AAR05707 standard; protein; 2324 AA.
XX
XX AC AAR05707;
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 16-AUG-1990 (first entry)
XX
XX DE Acetyl-CoA-carboxylase.
XX
XX KW Acetyl-CoA-carboxylase; Co2 ligase; malonyl-CoA; biotin;
XX biotinincarboxylase; carboxyl-transferase; acetyl-CoA; ATP; ds.
XX
XX OS Gallus sp.
XX
XX PN JP02057179-A.
XX
XX PD 26-FEB-1990.
XX
XX PF 24-AUG-1988; 88JP-00208170.
XX
XX PR 24-AUG-1988; 88JP-00208170.
XX
XX PA (MEIP) MEIJI MILK PROD CO LTD.
XX
XX DR WPI; 1990-104915/14.
XX
XX DR N-PSDB; AAQ04013.
XX
XX PT Acetyl-coa-carboxylase - derived from chicken liver, has known nucleotide
PT and amino acid sequence.

XX PS Disclosure; Fig 2; 14pp; Japanese.

XX CC Gene derived from chicken liver. (Updated on 25-MAR-2003 to correct PA field.)

XX CC

XX CC

SQ Sequence 2324 AA;

Query Match 73.5%; Score 36; DB 2; Length 2324;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

Db 1467 ESVRSWMRY 1476

RESULT 39

ABP70270

ID ABP70270 standard; protein; 2346 AA.

XX AC ABP70270;

XX DT 07-APR-2003 (first entry)

XX DE Amino acid sequence of human acetyl coenzyme A-carboxylase alpha.

XX KW Acetyl coenzyme A-carboxylase alpha; ACC-alpha; BRCA1; breast cancer; ovarian cancer.

XX KW

XX OS Homo sapiens.

XX PN WO2002100897-A2.

XX PD 19-DEC-2002.

XX PF 12-JUN-2002; 2002WO-FR002016.

XX PR 13-JUN-2001; 2001FR-00007740.

XX PR 05-MAR-2002; 2002FR-00002789.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.

XX PI Dalla Venezia N, Magnard C, Lenoir G, Sinilnikova-Brard O;

XX DR WPI; 2003-148772/14.

XX DR N-PSDB; AB223503.

XX PT New molecular complex of BRCA1 protein and acetyl coenzyme A-carboxylase alpha, useful in screening agents suitable for treatment, prevention or diagnosis of cancer.

XX PT

XX PS Disclosure; Page 55-62; 62pp; French.

XX CC The present sequence represents human acetyl coenzyme A-carboxylase alpha (ACC-alpha). The ACC-alpha protein was used in the course of the invention. The specification describes a molecular complex comprising a polypeptide having amino acids 1640-1663 of the human BRCA1 protein (or similar sequence from some other animal species), and a polypeptide that is a part of the acetyl coenzyme A-carboxylase alpha (ACC-alpha) protein able to bind the BRCA1 protein. The complex is implicated in predisposition to cancer of breast and ovary. It is used to screen for compounds that modulate interaction between BRCA1 and ACC-alpha, which are potentially useful for treatment, prevention and diagnosis of cancer, and to identify endogenous ligands. Modulated formation of the complex can be used for diagnosis of cancer. Antibodies directed against specific parts of human ACC-alpha are useful for localization of the complex in cells

XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

Db 1490 ESVRSWMRY 1499

RESULT 41

ABP59195

ID ABP59195 standard; protein; 2346 AA.

XX AC ABP59197;

XX DT 01-MAY-2003 (first entry)

XX DE Human acetyl-Coenzyme A-carboxylase-alpha variant #1.

XX KW Human; enzyme; acetyl-Coenzyme A-carboxylase-alpha; ACC-alpha; cancer; breast; ovary.

XX KW

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 1405

FT /note= "Leu substituted with Val"

XX PN WO2002100896-A2.

XX PD 19-DEC-2002.

XX PF 12-JUN-2002; 2002WO-FR002015.

XX PR 13-JUN-2001; 2001FR-00007740.

XX PR 05-MAR-2002; 2002FR-00002788.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.

XX PI Dalla Venezia NL, Magnard CM, Lenoir GM, Sinilnikova-Brard O;

XX DR WPI; 2003-175165/17.

XX PT In vitro diagnosis of cancer, particularly breast and ovarian cancer, or susceptibility, comprises detecting alterations in the acetyl coenzyme A-carboxylase alpha gene or protein expression.

XX PS Example 2; Page; 56pp; French.

XX CC The present sequence is a variant sequence for human acetyl-Coenzyme A-carboxylase-alpha (ACC-alpha). This sequence can be used for in vitro diagnosis of cancer (or of an increased risk of developing it), by detecting ACC-alpha gene mutations or polymorphisms, or altered ACC-alpha protein expression, relative to a control population. The method is particularly used to diagnose cancer, especially of breast or ovary, or for assessing the risk of developing such cancers. Note: The present sequence was not shown in the specification, but was derived from information given

XX SQ Sequence 2346 AA;

Query Match 73.5%; Score 36; DB 6; Length 2346;

Best Local Similarity 70.0%; Pred. No. 1.3e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

Db 1490 ESVRSWMRY 1499

XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX Tong L, Zhang H, Yang Z;
 XX WPI; 2004-571486/55.
 XX New crystallizable composition comprising a carboxyltransferase domain of
 PT acetyl-coenzyme A (CoA) carboxylase, for treating e.g. diabetes, obesity,
 PT cardiovascular disease, atherosclerosis, or cancer.
 XX Example 4; SEQ ID NO 7; 195pp; English.
 XX The invention comprises a crystallisable composition containing a
 CC carboxyltransferase (CT) domain of acetyl-coenzyme A (acetyl-CoA)
 CC carboxylase. The composition of the invention is useful for treating:
 CC metabolic syndrome, diabetes, obesity, cardiovascular disease,
 CC atherosclerosis, depression, cancer, hyperlipidaemia, dyslipidaemia,
 CC hypertension, hyperuricaemia and renal dysfunction. The present amino
 CC acid sequence represents a human acetyl-CoA carboxylase of the invention.
 XX Sequence 2346 AA;
 SQ
 Query Match 73.5%; Score 36; DB 8; Length 2346;
 Best Local Similarity 70.0%; Pred. No. 1.3e+03;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ESVREQVMKY 10
 DB 1490 ESVRSVMRY 1499
 ||||| ||:
 RESULT 44
 ADA33386
 ID ADA33386 standard; protein; 232 AA.
 XX
 AC ADA33386;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #547.
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 DR N-PSDB; ADA29260.
 XX
 XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 XX Example; SEQ ID NO 4673; 328pp; English.
 XX
 XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX Sequence 232 AA;
 SQ
 Query Match 71.4%; Score 35; DB 6; Length 232;
 Best Local Similarity 60.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ESVREQVMKY 10
 DB 8 ESLSEQIVKY 17
 ||: ||:
 RESULT 45
 ABR53195
 ID ABR53195 standard; protein; 810 AA.
 XX
 AC ABR53195;
 XX
 DT 20-JUN-2003 (first entry)
 XX
 DE Protein sequence #SEQ ID 1255.
 XX
 KW Multiprotein complex; eukaryote; drug target; diagnosis.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN EP1258494-A1.
 XX
 PD 20-NOV-2002.
 XX
 PF 20-DEC-2001; 2001EP-00130253.
 XX
 PR 15-MAY-2001; 2001EP-00111774.
 XX
 PA (CELL-) CELLZONE AG.
 XX
 PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
 PI Marzioch M, Schultz JD, Superti-Furga GD;
 XX
 DR WPI; 2003-250078/25.
 DR N-PSDB; ACC61237.
 XX
 XX New isolated protein complexes useful for diagnosing a disease or
 PT disorder, or as a target for an active agent of a pharmaceutical,
 PT preferably a drug target in the treatment or prevention of disease or
 PT disorder.
 XX
 PS Disclosure; SEQ ID NO 1255; 17pp + Sequence Listing; English.
 XX
 XX The invention relates to multiprotein complexes from eukaryotes. Proteins
 CC of the invention and DNA sequences encoding them are given in records
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
 CC obtainable by using a protein as a bait and isolating the set of proteins
 CC which is attached thereto from cells. Such protein complexes may comprise
 CC up to 30 distinct proteins. Protein complexes of the invention are useful
 CC for diagnosing a disease or disorder, or as a target for an active agent
 CC of a pharmaceutical, preferably a drug target in the treatment or
 CC prevention of a disease or disorder. Note: The sequence data for this
 CC patent is not represented in the printed specification, but is based on
 CC sequence information supplied by the European Patent Office. The complete
 CC document is available on CD-ROM
 XX
 XX Sequence 810 AA;
 SQ
 Query Match 71.4%; Score 35; DB 6; Length 810;
 Best Local Similarity 70.0%; Pred. No. 6.4e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 |||||:|:
 Db 13 ESVBEAVLKY 22

RESULT 46

ADK62816
 ID ADK62816 standard; protein; 810 AA.

XX AC ADK62816;

DT 06-MAY-2004 (first entry)

DE Disease treating protein complex-derived protein #534.

XX protein complex; drug target; diagnosis.

XX Unidentified.

XX EPI338608-A2.

PD 27-AUG-2003.

PF 20-DEC-2002; 2002EP-00102902.

XX 20-DEC-2001; 2001EP-00130253.

XX (CELL-) CELLZONE AG.

XX Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;

PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;

PI Michon A, Leutwein C, Rick J;

XX WPI; 2003-638460/61.

DR N-PSDB; ADK62817.

XX New proteins and protein complexes from eukaryotes, useful as targets in
 PT drug screening, or in diagnosing or screening for the presence of a
 PT disease or disorder, or a predisposition for developing a disease or
 PT disorder in a subject.

XX Disclosure; SEQ ID NO 1067; 13pp; English.

XX The invention relates to novel protein complexes comprising a first and a
 CC second protein, or its derivative, fragment, homologue or variant. The
 CC proteins are selected from given protein complexes, which are not defined
 CC in the specification. The variants are encoded by nucleic acids that
 CC hybridize to the nucleic acids encoding the proteins under low stringency
 CC conditions. The protein complexes are useful as targets for an active
 CC agent of a pharmaceutical. These protein complexes are particularly
 CC useful as drugs targets for the treatment or preventing of a disease or
 CC disorder. The complexes and methods above are useful in diagnosing or
 CC screening for the presence of a disease or disorder or a predisposition
 CC for developing a disease or disorder in a subject. These are also useful
 CC in screening for a drug for treatment or prevention of a disease or
 CC disorder. The molecule that modulates the amount, activity or protein
 CC components of the complex is useful for the manufacture of a medicament
 CC for the treatment or prevention of a disease or disorder. This sequence
 CC corresponds to a protein of the invention. (Note: the sequence data for
 CC this patent did not form part of the printed specification but was
 CC obtained from the EPO in electronic format).

XX Sequence 810 AA;

Query Match 71.4%; Score 35; DB 7; Length 810;
 Best Local Similarity 70.0%; Pred. No. 6.4e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 |||||:|:
 Db 13 ESVBEAVLKY 22

RESULT 47

ADO16901

ID ADO16901 standard; protein; 1245 AA.

XX AC ADO16901;

DT 12-AUG-2004 (first entry)

DE Tobacco budworm voltage gated calcium channel subunit protein SeqID 14.

XX voltage-gated calcium channel; pesticide; agrochemical; veterinary;
 KW pharmaceutical; insecticide; nematocide; molluscicide; helminthicide;
 KW acaricide; biocide; crop protection agent; antiparasitic;
 KW tobacco budworm.

XX Heliothis virescens.

PN WO2004044553-A2.

XX 27-MAY-2004.

XX 10-NOV-2003; 2003WO-US036001.

XX 08-NOV-2002; 2002US-0425077P.

XX (FWCC) FMC CORP.

XX Wu S, Hayashi JH, Kinne LP, Dierks PM;

XX WPI; 2004-411759/38.

DR N-PSDB; ADO16900.

XX New Lepidoptera calcium channel subunits, useful as active agents in
 PT agrochemical, veterinary or pharmaceutical fields, e.g. as insecticide,
 PT acaricide, or for treating infection, damage or discomfort caused by
 PT parasitic organisms.

XX Claim 1; SEQ ID NO 14; 144pp; English.

XX This invention relates to novel isolated nucleic acids from Lepidoptera
 CC that each encode a voltage-gated calcium channel subunit protein.
 CC Specifically, it refers to methods and screening assays to identify
 CC compounds that can target and modulate these channels and as such that
 CC have potential activity as pesticides. Accordingly, the present invention
 CC provides active agents that have applications in the agrochemical,
 CC veterinary and pharmaceutical fields, and can be used as an insecticide,
 CC nematocide, molluscicide, helminthicide, acaricide or other types of
 CC pesticides or biocides. In addition, they can be applied systemically (as
 CC additives) to the soil or habitat of the organism to act as crop
 CC protection agents or as a pesticide for household use to prevent or treat
 CC infection and/ or discomfort caused by parasitic organisms such as
 CC parasitic arthropods, nematodes or helminths. This polypeptide sequence
 CC is a tobacco budworm voltage-gated calcium channel protein subunit of the
 CC invention.

XX Sequence 1245 AA;

Query Match 71.4%; Score 35; DB 8; Length 1245;
 Best Local Similarity 70.0%; Pred. No. 1e+03;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10

Db 415 ESVREQVMKY 424

RESULT 48

AAY22382

ID AAY22382 standard; peptide; 15 AA.

XX AC AAY22382;

DT 27-SEP-1999 (first entry)

PT expression profile or pattern of a gene and identifying homologous gene.

XX PS Claim 17; SEQ ID NO 6582; 246pp + Sequence Listing; English.

XX CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids.
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX SQ Sequence 212 AA;

Query Match 69.4%; Score 34; DB 4; Length 212;
Best Local Similarity 60.0%; Pred. NO. 2.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 128 BEIREQVLLY 137

Search completed: May 12, 2006, 10:36:36
Job time : 102.75 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds

(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-38

Perfect score: 49

Sequence: 1 ESVREQVMKY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : PIR 80:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	81.6	1164	2 G71827	hypothetical prote
2	38	77.6	291	2 S73600	probable GTP-bind
3	37	75.5	429	2 A47305	translation initia
4	36	73.5	151	2 S54048	ribosomal protein
5	36	73.5	160	2 F64816	ybiA protein - Esc
6	36	73.5	1234	2 T30254	jumonji protein -
7	36	73.5	2324	1 A29924	acetyl-CoA carboxy
8	36	73.5	2345	1 A35578	acetyl-CoA carboxy
9	36	73.5	2346	2 I38928	acetyl-CoA carboxy
10	35	71.4	486	2 G72358	conserved hypothet
11	35	71.4	810	1 S45307	myb-related protei
12	35	71.4	1279	1 E69681	peptide synthetase
13	34	69.4	212	2 G82176	probable outer mem
14	34	69.4	272	2 B90544	hypothetical prote
15	34	69.4	357	2 AD1062	protein kinase [im
16	34	69.4	455	2 E90514	hypothetical prote
17	34	69.4	595	2 T17590	probable glutamine
18	34	69.4	786	2 S71091	acetyl-CoA carboxy
19	34	69.4	825	2 S54465	YTA12 protein prec
20	34	69.4	916	2 C38992	cadherin 4 precurs
21	34	69.4	2339	2 S41121	acetyl-CoA carboxy
22	33	67.3	179	2 B90408	hypothetical prote
23	33	67.3	243	2 T25190	hypothetical prote
24	33	67.3	278	2 T03942	rRNA N-glycosidase
25	33	67.3	282	2 A84341	hypothetical prote
26	33	67.3	298	2 G82581	GTP binding protei
27	33	67.3	355	2 H69486	conserved hypothet
28	33	67.3	358	2 T43289	fructose-bisphosph
29	33	67.3	358	2 T39798	fructose-bisphosph

30	67.3	392	2	AE2092	hypothetical prote
31	67.3	431	2	T12450	hypothetical prote
32	67.3	490	2	F86841	iron-binding oxida
33	67.3	514	2	E84667	probable cytochrom
34	67.3	570	2	A45249	alpha-glucosidase
35	67.3	572	2	T41371	hypothetical prote
36	67.3	595	1	A48077	myb-related protei
37	67.3	1153	2	A97179	ATP-dependent exon
38	67.3	1202	2	T23429	hypothetical prote
39	67.3	1873	2	A30063	dihydropyridine re
40	65.3	92	2	S69116	fibronogen gamma c
41	65.3	138	2	T27724	hypothetical prote
42	65.3	217	2	E97310	uncharacterized co
43	65.3	222	2	H64321	hypothetical prote
44	65.3	301	2	S45880	SCO1 protein homol
45	65.3	314	2	S00909	replication initia
46	65.3	314	2	A30480	repJ protein - Sta
47	65.3	316	2	T49224	SRG1-like protein
48	65.3	319	2	T04309	acidic ribosomal p
49	65.3	319	2	T03944	acidic ribosomal p
50	65.3	350	2	T22701	hypothetical prote
51	65.3	376	2	S49801	OGG1 protein - yea
52	65.3	403	2	E82810	methionine adenosy
53	65.3	403	2	D96961	carboxyl-terminal
54	65.3	410	2	G72586	hypothetical prote
55	65.3	444	2	S05313	fibronogen gamma-B
56	65.3	445	2	E83805	hypothetical prote
57	65.3	445	2	H97075	4-aminobutyrate am
58	65.3	489	1	D64311	fumarate reductase
59	65.3	501	2	G72368	comM protein - The
60	65.3	530	2	T46648	zinc finger protei
61	65.3	550	2	H84830	hypothetical prote
62	65.3	570	1	S50933	myb-related protei
63	65.3	678	2	T50256	probable vacuolar
64	65.3	721	2	H82528	L-ascorbate oxidas
65	65.3	759	2	B81101	ribonucleoside-dip
66	65.3	782	2	AD0119	conserved hypothet
67	65.3	782	2	AF0179	cadherin 8 - human
68	65.3	793	2	D38992	hypothetical prote
69	65.3	845	2	T00071	preprotein translo
70	65.3	930	2	AC2412	exonuclease ABC c
71	65.3	940	2	A82329	hypothetical prote
72	65.3	1059	2	T22545	genome polyprotein
73	65.3	2156	1	RRVUNE	lsu ribosomal prot
74	63.3	91	2	G75094	probable ribosomal
75	63.3	94	2	E71071	uncharacterized co
76	63.3	96	2	F97355	30S ribosomal prot
77	63.3	120	2	E81229	hypothetical prote
78	63.3	127	2	C71713	absicisic acid-indu
79	63.3	130	2	A48892	hypothetical prote
80	63.3	139	2	E97709	hypothetical prote
81	63.3	166	2	G83399	hypothetical prote
82	63.3	188	2	G96652	protein F23N19.15
83	63.3	189	2	C81428	pepidyl-prolyl ci
84	63.3	204	1	MNNZSV	nonstructural prot
85	63.3	204	1	MNNZSV	nonstructural prot
86	63.3	206	2	A98106	hypothetical prote
87	63.3	206	2	F95241	hydrolase, haloaci
88	63.3	259	2	S36410	translation initia
89	63.3	268	2	E82472	hypothetical prote
90	63.3	283	2	A96655	hypothetical prote
91	63.3	291	2	A75096	UTP-glucose-1-phos
92	63.3	292	2	B90454	dehydrogenase [imp
93	63.3	292	2	AH1354	ABC transporter (A
94	63.3	294	2	S47535	homeodomain-contai
95	63.3	295	2	S29045	estrone sulfotrans
96	63.3	311	1	RQ9AD2	repB protein - Sta
97	63.3	314	2	A26473	replication protei
98	63.3	316	1	A39763	aldehyde reductase
99	63.3	330	2	B97886	hypothetical prote
100	63.3	341	2	B86598	flagellar M-ring p
101	63.3	341	2	D72025	flagellar m-ring p
102	63.3	386	2	C70505	probable acyl-coa

103	31	63.3	415	2	A10361	conserved hypother	176	30	61.2	553	2	AC2931	hypothetical prote
104	31	63.3	430	2	S62500	actin-related prot	177	30	61.2	603	2	H69121	hypothetical prote
105	31	63.3	443	2	JC4067	hypothetical 50.1k	178	30	61.2	619	2	S30780	hexose metabolism-
106	31	63.3	488	2	A95926	probable argininos	179	30	61.2	626	2	D88601	protein Y49E10.11
107	31	63.3	590	2	A82350	conserved hypother	180	30	61.2	630	2	C98351	oligopeptide-bindi
108	31	63.3	592	2	AD2328	hypothetical prote	181	30	61.2	657	2	G86590	exinuclease ABC su
109	31	63.3	632	2	H69590	asparagine synthet	182	30	61.2	657	2	B72034	exinuclease ABC,
110	31	63.3	765	2	E88924	protein R02C2.3 [i	183	30	61.2	698	2	G89787	hypothetical prote
111	31	63.3	856	2	E70875	hypothetical prote	184	30	61.2	727	2	T26096	hypothetical prote
112	31	63.3	924	2	C86725	chromosome segrega	185	30	61.2	752	2	G96510	hypothetical prote
113	31	63.3	925	2	T33732	probable excinucle	186	30	61.2	759	2	G81841	ribonucleoside-dip
114	31	63.3	932	2	C70191	penicillin-binding	187	30	61.2	769	1	ISBYT1	DNA topoisomerase
115	31	63.3	1043	2	AH1906	hypothetical prote	188	30	61.2	774	1	I40421	endopeptidase La (
116	31	63.3	1076	2	T38122	hypothetical prote	189	30	61.2	774	2	B84031	ATP-dependent prot
117	31	63.3	1116	2	T14598	polypeptin - slim	190	30	61.2	779	2	T05990	hypothetical prote
118	31	63.3	1164	2	T44806	hypothetical prote	191	30	61.2	783	2	I50116	N-cadherin precurs
119	31	63.3	2223	2	A47447	calcium channel pr	192	30	61.2	783	2	AH3592	ribonuclease R (EC
120	31	61.2	87	2	T10462	hypothetical prote	193	30	61.2	786	2	F84608	hypothetical prote
121	30	61.2	110	2	S74013	hypothetical prote	194	30	61.2	790	2	G02678	cadherin-14 - huma
122	30	61.2	115	2	E36811	hypothetical prote	195	30	61.2	791	2	D82901	ATP-dependent prot
123	30	61.2	118	2	AH0028	30S ribosomal prot	196	30	61.2	796	2	I49556	cadherin-11 - mous
124	30	61.2	134	2	I47010	gastrin-releasing	197	30	61.2	798	2	G83420	Lon peptidase PAL
125	30	61.2	140	2	F75335	probable transposa	198	30	61.2	799	2	C87492	ATP-dependent prot
126	30	61.2	144	2	A12573	hypothetical prote	199	30	61.2	805	2	A75014	hypothetical prote
127	30	61.2	151	1	S25374	ribosomal protein	200	30	61.2	807	2	T00990	hypothetical prote
128	30	61.2	151	1	S26296	40S ribosomal prot	201	30	61.2	810	1	JC6045	endopeptidase La (
129	30	61.2	154	2	D89137	hypothetical prote	202	30	61.2	820	2	H81106	ATP-dependent prot
130	30	61.2	194	2	G82091	RNA polymerase sig	203	30	61.2	820	2	H81908	probable endopepti
131	30	61.2	201	2	A10928	probable phage tai	204	30	61.2	858	2	D71711	endopeptidase Clp
132	30	61.2	201	2	A11035	probable phage tai	205	30	61.2	877	1	IJBOCN	N-cadherin precurs
133	30	61.2	210	2	F96551	hypothetical prote	206	30	61.2	881	1	B71316	endopeptidase La (
134	30	61.2	226	2	AC0176	probable exported	207	30	61.2	889	2	T20123	hypothetical prote
135	30	61.2	229	2	A99762	hypothetical prote	208	30	61.2	896	2	S76064	hypothetical prote
136	30	61.2	241	1	BVECPU	peripheral membran	209	30	61.2	896	2	S59990	phycobillosome anch
137	30	61.2	241	2	AF0956	phosphate transpor	210	30	61.2	905	1	IJXLC1	N-cadherin 1 precu
138	30	61.2	241	2	D86057	hypothetical prote	211	30	61.2	905	1	RGBYS5	regulatory protein
139	30	61.2	241	2	D91211	hypothetical prote	212	30	61.2	906	1	IJHUCN	cadherin 2 precurs
140	30	61.2	249	2	F95957	probable transcrip	213	30	61.2	906	1	IJMSCN	N-cadherin precurs
141	30	61.2	249	2	E84072	hypothetical prote	214	30	61.2	906	1	IJXLC2	N-cadherin 2 precu
142	30	61.2	256	2	A75005	hypothetical prote	215	30	61.2	906	2	T28034	hypothetical prote
143	30	61.2	256	2	H71157	hypothetical prote	216	30	61.2	912	1	IJHCHN	N-cadherin precurs
144	30	61.2	262	2	G82280	vibriobactin-speci	217	30	61.2	913	1	A47543	N-cadherin precurs
145	30	61.2	273	2	F71318	conserved hypother	218	30	61.2	913	1	IJHCHR	N-cadherin precurs
146	30	61.2	284	2	H86183	protein T7A14.13 [219	30	61.2	976	2	S45738	pleiotropic drug r
147	30	61.2	295	2	F59102	probable UTP-gluc	220	30	61.2	1007	2	S48535	rho-type GFraser-ac
148	30	61.2	303	1	CFCYB	carboxypeptidase B	221	30	61.2	1080	2	A35088	phycobillosome link
149	30	61.2	305	2	T34177	hypothetical prote	222	30	61.2	1123	2	D96756	receptor-like prot
150	30	61.2	321	1	R5UBPO	acidic ribosomal p	223	30	61.2	1132	2	AD1809	phycobillosome core
151	30	61.2	335	2	T25110	hypothetical prote	224	30	61.2	1286	2	A42150	P-glycoprotein pgp
152	30	61.2	345	2	F87844	protein T22C1.6 [i	225	30	61.2	1314	2	S19488	probable membrane
153	30	61.2	367	2	H82437	2-aminoethylphosph	226	30	61.2	1802	2	H88444	protein C26B6.12 [
154	30	61.2	382	2	C87473	acyl-CoA dehydroge	227	30	61.2	1811	2	T00035	nonstructural poly
155	30	61.2	387	2	F74022	hypothetical prote	228	30	61.2	2151	1	S16449	genome polyprotein
156	30	61.2	396	2	S84326	GTP-binding protei	229	30	61.2	2160	2	T20241	hypothetical prote
157	30	61.2	406	1	S48220	serine-type D-Ala-	230	30	61.2	2206	2	JCS280	voltage-dependent
158	30	61.2	409	2	T14611	alanine transamina	231	30	61.2	2523	2	T18477	hypothetical prote
159	30	61.2	430	2	G86870	histidine-trNA lig	232	29	59.2	76	2	AG0683	hypothetical prote
160	30	61.2	432	2	E89954	conserved hypother	233	29	59.2	82	2	AB2351	hypothetical prote
161	30	61.2	448	2	T18710	hypothetical prote	234	29	59.2	83	1	W8BPT3	gene 18.7 protein
162	30	61.2	459	2	JC9631	cellobiohydrolase	235	29	59.2	97	2	S06782	tryptophan synthas
163	30	61.2	463	2	S06309	vicilin precursor	236	29	59.2	98	2	AC1033	hypothetical prote
164	30	61.2	463	2	A27288	vicilin precursor	237	29	59.2	134	2	E57233	complexin II - hum
165	30	61.2	466	2	E75201	pyridoxal phosphat	238	29	59.2	134	2	C57233	synaphin - bovine
166	30	61.2	470	2	F72238	conserved hypother	239	29	59.2	134	2	JC4226	complexin II - mou
167	30	61.2	471	2	A45616	antigenic protein	240	29	59.2	134	2	D57233	hypothetical prote
168	30	61.2	481	2	A30203	prolyl -tRNA synth	241	29	59.2	136	2	B97881	H+-transporting tw
169	30	61.2	483	2	AB0072	tagaturonate reduc	242	29	59.2	156	2	C81025	H+-transporting tw
170	30	61.2	489	2	C84401	aldehyde dehydroge	243	29	59.2	157	2	AG3560	transcription regu
171	30	61.2	497	2	T14609	alanine transamina	244	29	59.2	158	2	A84069	hypothetical prote
172	30	61.2	501	2	T37747	hypothetical prote	245	29	59.2	165	2	AC0451	conserved hypochet
173	30	61.2	506	1	PWEGA	H+-transporting tw	246	29	59.2	172	2	B64320	acetolactate synth
174	30	61.2	518	2	B84514	probable cytochrom	247	29	59.2	180	2	F69408	phosphoribosylamin
175	30	61.2	532	2	B69343	2-oxoacid-ferredox	248	29	59.2	181	2	D81285	hypothetical prote

249	29	59.2	181	2	T44572	322	29	59.2	531	2	T25537	hypothetical prote
250	29	59.2	181	2	D75052	323	29	59.2	532	2	B64100	purH bifunctional
251	29	59.2	181	2	A71108	324	29	59.2	545	2	F84533	Mutator-like trans
252	29	59.2	196	2	AF0699	325	29	59.2	564	2	AD1550	heat shock protein
253	29	59.2	197	2	F63359	326	29	59.2	571	2	C84038	DNA-dependent DNA
254	29	59.2	198	2	S73275	327	29	59.2	588	2	C95252	L-fucose isomerase
255	29	59.2	224	2	H84829	328	29	59.2	588	2	A99717	probable long-chain
256	29	59.2	223	2	C86849	329	29	59.2	598	2	T35513	conserved hypothe
257	29	59.2	235	2	E72217	330	29	59.2	611	2	D72275	conserved hypothe
258	29	59.2	248	2	T23708	331	29	59.2	648	2	H84587	hypothetical prote
259	29	59.2	248	2	C84741	332	29	59.2	650	2	H81708	hypothetical prote
260	29	59.2	254	2	B73366	333	29	59.2	650	2	AG0733	probable bacterio
261	29	59.2	261	2	T46138	334	29	59.2	651	2	AH3276	lysostaphin (EC 3
262	29	59.2	267	2	D70376	335	29	59.2	667	2	B96575	hypothetical prote
263	29	59.2	274	2	A11256	336	29	59.2	677	2	F95232	hypothetical prote
264	29	59.2	274	2	A16119	337	29	59.2	680	2	H98096	immunity protein,
265	29	59.2	275	2	G89961	338	29	59.2	686	2	C64428	conserved hypothe
266	29	59.2	287	2	G70974	339	29	59.2	702	2	T03903	hypothetical prote
267	29	59.2	287	2	A70614	340	29	59.2	711	2	S67293	probable membrane
268	29	59.2	290	2	D81359	341	29	59.2	726	2	AG0837	TonB-dependent out
269	29	59.2	290	2	H64242	342	29	59.2	746	2	A75018	transport protein
270	29	59.2	296	1	F69257	343	29	59.2	759	2	G69258	ATP-dependent RNA
271	29	59.2	310	2	T01224	344	29	59.2	762	2	T00410	protein kinase hom
272	29	59.2	310	2	F64311	345	29	59.2	779	2	E97778	endopeptidase La (
273	29	59.2	312	2	E86705	346	29	59.2	784	1	A71704	ATP-dependent LA p
274	29	59.2	317	2	F96810	347	29	59.2	786	2	G82141	ribonucleoside-dip
275	29	59.2	323	2	G86877	348	29	59.2	785	2	A69927	ribonucleoside-dip
276	29	59.2	325	2	A82073	349	29	59.2	789	2	T42564	cadherin 11 precu
277	29	59.2	335	2	AC0786	350	29	59.2	796	2	A38952	cadherin-11 - mous
278	29	59.2	335	2	A64988	351	29	59.2	796	2	I48277	OB-cadherin precu
279	29	59.2	335	2	H85857	352	29	59.2	796	2	A53584	hypothetical prote
280	29	59.2	335	2	F91013	353	29	59.2	804	2	T18014	hypothetical prote
281	29	59.2	338	2	G83508	354	29	59.2	807	2	D59102	collagenase - Meth
282	29	59.2	339	2	T22571	355	29	59.2	817	2	F86742	ribonuclease (limp
283	29	59.2	341	2	F87620	356	29	59.2	864	2	A84974	translation initia
284	29	59.2	345	1	MFVNSY	357	29	59.2	864	2	D97351	sensor protein kdp
285	29	59.2	355	2	T02056	358	29	59.2	900	2	T00039	hypothetical prote
286	29	59.2	367	2	T29752	359	29	59.2	906	2	T38371	probable helicase
287	29	59.2	370	2	D84464	360	29	59.2	922	2	B38371	excinuclease ABC c
288	29	59.2	377	2	A53384	361	29	59.2	948	2	B81883	sensor protein Rcs
289	29	59.2	377	2	S60758	362	29	59.2	948	2	AD0790	excinuclease ABC c
290	29	59.2	378	2	JH0134	363	29	59.2	965	2	C82560	excinuclease ABC c
291	29	59.2	387	1	A56466	364	29	59.2	965	2	C82560	probable zinc-bind
292	29	59.2	390	2	B95954	365	29	59.2	999	2	T36021	DNA topoisomerase
293	29	59.2	392	2	I51422	366	29	59.2	1003	2	T13951	SEC14 protein homo
294	29	59.2	392	2	S32405	367	29	59.2	1008	2	T41244	hypothetical prote
295	29	59.2	392	2	T46418	368	29	59.2	1013	1	TQECT	transposase - Esch
296	29	59.2	396	2	A47151	369	29	59.2	1015	1	TQECT	fibronectin-bindin
297	29	59.2	397	2	S06114	370	29	59.2	1018	2	A32192	RNA helicase HEL11
298	29	59.2	425	2	D72345	371	29	59.2	1032	2	A57514	beta-galactosidase
299	29	59.2	432	2	T05236	372	29	59.2	1034	2	T30551	beta-galactosidase
300	29	59.2	434	2	A43252	373	29	59.2	1034	2	T30574	beta-galactosidase
301	29	59.2	436	2	C95222	374	29	59.2	1078	2	AF3486	DNA-directed DNA p
302	29	59.2	442	2	B82899	375	29	59.2	1083	2	A77647	DNA polymerase III
303	29	59.2	444	2	H75544	376	29	59.2	1084	2	T12925	ribonucleotide red
304	29	59.2	451	2	C98086	377	29	59.2	1134	2	D75014	hypothetical prote
305	29	59.2	452	2	AB2965	378	29	59.2	1145	2	T18235	transcription acti
306	29	59.2	452	2	G90654	379	29	59.2	1199	2	T37561	probable transcrip
307	29	59.2	452	2	G85505	380	29	59.2	1224	2	S73171	DNA-directed RNA p
308	29	59.2	452	2	E64745	381	29	59.2	1224	2	T47993	hypothetical prote
309	29	59.2	455	2	S50391	382	29	59.2	1232	2	T24897	hypothetical prote
310	29	59.2	455	2	A70461	383	29	59.2	1254	2	F64237	DNA-directed RNA p
311	29	59.2	465	1	F64630	384	29	59.2	1292	2	A28313	glued protein - fr
312	29	59.2	474	2	AC1904	385	29	59.2	1319	2	A28313	probable xanthine
313	29	59.2	475	2	T35799	386	29	59.2	1364	2	T51920	myosin-2 isoform -
314	29	59.2	484	2	G64398	387	29	59.2	1471	2	T40117	calcium channel, v
315	29	59.2	491	2	B86155	388	29	59.2	1873	2	A55645	vitellogenin I pre
316	29	59.2	492	1	S28003	389	29	59.2	1912	2	T29088	adherence factor T
317	29	59.2	497	2	G98318	390	29	59.2	3255	2	G81702	hypothetical prote
318	29	59.2	498	2	F84667	391	29	59.2	88	2	T17418	hypothetical prote
319	29	59.2	499	2	T32688	392	29	57.1	91	2	T16579	hypothetical prote
320	29	59.2	503	2	AB1933	393	29	57.1	100	4	S32195	hypothetical prote
321	29	59.2	527	2	T26732	394	29	57.1	112	2	F70403	hypothetical prote

395	28	57.1	113	2	G69065	hypothetical prote	468	28	57.1	301	2	E85902	GTP-binding protei
396	28	57.1	115	2	A83554	hypothetical prote	469	28	57.1	301	2	AB0829	GTP-binding protei
397	28	57.1	122	2	B64493	hypothetical prote	470	28	57.1	301	2	AF1257	GTP-binding protei
398	28	57.1	124	2	AC0803	probable bacteriop	471	28	57.1	301	2	AB1620	GTP-binding protei
399	28	57.1	131	2	AG1082	hypothetical prote	472	28	57.1	301	2	H91057	GTP-binding protei
400	28	57.1	131	2	AH1439	hypothetical prote	473	28	57.1	302	2	A34406	aldelyde reductase
401	28	57.1	134	2	S66294	921-S protein - mo	474	28	57.1	302	2	F64042	GTP-binding protei
402	28	57.1	134	2	A57233	complexin I - rat	475	28	57.1	303	2	B69584	3-methyladenine DN
403	28	57.1	136	2	F70182	hypothetical prote	476	28	57.1	303	2	S45093	Na+/K+-exchanging
404	28	57.1	147	2	F70884	hypothetical prote	477	28	57.1	303	2	A72365	conserved hypothet
405	28	57.1	147	2	T26225	hypothetical prote	478	28	57.1	303	2	AG0331	probable GTP-bindi
406	28	57.1	148	2	F75091	hypothetical prote	479	28	57.1	305	2	F83548	GTP-binding protei
407	28	57.1	151	2	A84163	hypothetical prote	480	28	57.1	305	2	F69306	hypothetical prote
408	28	57.1	157	2	H75178	hypothetical prote	481	28	57.1	307	2	D81935	probable K04P1.15
409	28	57.1	159	2	B84000	hypothetical prote	482	28	57.1	309	2	D88956	protein K04P1.15
410	28	57.1	162	2	B81967	probable periplasm	483	28	57.1	309	2	H71089	hypothetical prote
411	28	57.1	162	2	A70489	conserved hypothet	484	28	57.1	311	1	IQBS44	primosome componen
412	28	57.1	165	2	A10843	conserved hypothet	485	28	57.1	311	2	C75493	GTP-binding protei
413	28	57.1	166	2	G81021	thioredoxin, proba	486	28	57.1	312	2	S61213	malate dehydrogena
414	28	57.1	172	2	H72742	hypothetical prote	487	28	57.1	314	2	JT0372	repi protein - Sta
415	28	57.1	173	2	A81237	conserved hypothet	488	28	57.1	316	1	A60603	aldehyde reductase
416	28	57.1	173	2	G84435	hypothetical prote	489	28	57.1	316	2	I49484	aldehyde reductase
417	28	57.1	183	2	G70450	hypothetical prote	490	28	57.1	318	2	C72856	AcOrf-51 protein -
418	28	57.1	185	2	E95904	hypothetical prote	491	28	57.1	319	2	F85054	probable malonyl-C
419	28	57.1	190	2	AH2529	transcription regu	492	28	57.1	320	2	E91022	o-succinylbenzoyl-
420	28	57.1	191	2	B82715	chorismate mutase	493	28	57.1	320	2	F85866	hypothetical prote
421	28	57.1	196	2	G83541	probable transcrip	494	28	57.1	320	2	C64997	o-succinylbenzoate
422	28	57.1	199	1	S25967	NADH2 dehydrogenas	495	28	57.1	323	2	T22492	hypothetical prote
423	28	57.1	199	1	TVMSB1	transforming prote	496	28	57.1	324	2	F81169	GTP-binding protei
424	28	57.1	205	1	TVHUB1	transforming prote	497	28	57.1	326	2	C64937	hypothetical prote
425	28	57.1	207	2	A69349	conserved hypothet	498	28	57.1	326	2	H90938	hypothetical prote
426	28	57.1	207	2	S56209	probable membrane	499	28	57.1	326	2	D85787	hypothetical prote
427	28	57.1	208	1	D71802	hypothetical prote	500	28	57.1	326	2	E98077	UDPglucose 4-epime
428	28	57.1	210	2	A40979	temperature shock-	501	28	57.1	327	2	A70435	cystine synthase
429	28	57.1	213	2	A83837	siroheme synthase	502	28	57.1	327	2	D84806	hypothetical prote
430	28	57.1	222	2	C97164	uncharacterized pr	503	28	57.1	331	1	A70032	conserved hypothet
431	28	57.1	236	1	TVMSA1	transforming prote	504	28	57.1	331	2	T14544	fructokinase (EC 2
432	28	57.1	236	2	JC7393	B-cell lymphoma 2	505	28	57.1	335	2	H75029	peptidase homolog
433	28	57.1	236	2	I67432	BCL-2 - rat (fragm	506	28	57.1	335	2	AD1314	hypothetical prote
434	28	57.1	236	2	I53744	gene bcl-2 protein	507	28	57.1	335	2	AH3298	sensory transducti
435	28	57.1	236	2	A82287	phosphate transpor	508	28	57.1	336	2	D95213	UDP-glucose 4-epim
436	28	57.1	238	2	E97105	probable membrane	509	28	57.1	336	2	T14437	inner envelope mem
437	28	57.1	239	1	TVHUA1	transforming prote	510	28	57.1	336	2	H82040	general secretion
438	28	57.1	243	2	S76271	hypothetical prote	511	28	57.1	339	2	E82211	conserved hypothet
439	28	57.1	246	2	B43942	myogenic regulator	512	28	57.1	346	2	F75377	ABC transporter, A
440	28	57.1	255	1	E71090	probable lactam ut	513	28	57.1	351	2	S21345	retrovirus-related
441	28	57.1	255	1	S04105	myogenic factor 5	514	28	57.1	357	2	G70010	probable aspartate
442	28	57.1	255	2	JN0624	Myogenic factor -	515	28	57.1	357	2	A83663	hypothetical prote
443	28	57.1	255	2	S22825	myf-5 protein - mo	516	28	57.1	359	2	T52337	phosphoprotein pho
444	28	57.1	256	1	B83852	phosphoesterase-re	517	28	57.1	359	2	F84513	hypothetical prote
445	28	57.1	257	2	E87290	transcription regu	518	28	57.1	360	2	AG1975	hypothetical prote
446	28	57.1	258	2	AC2575	chromosome partiti	519	28	57.1	362	2	D84952	GTP-binding protei
447	28	57.1	258	2	S41126	Myf5 protein - chi	520	28	57.1	370	2	H90559	hypothetical prote
448	28	57.1	259	2	A86822	hypothetical prote	521	28	57.1	379	2	H75318	membrane lipoprote
449	28	57.1	261	2	A95048	enoyl-CoA hydratase	522	28	57.1	379	2	A57477	potassium channel
450	28	57.1	261	2	G97918	enoyl-CoA hydratase	523	28	57.1	381	2	T24256	hypothetical prote
451	28	57.1	263	2	D90231	conserved hypothet	524	28	57.1	387	1	D84508	probable phosphoes
452	28	57.1	265	2	AD0620	probable prophage	525	28	57.1	387	2	T52650	cystine synthase
453	28	57.1	265	2	T40757	conserved hypothet	526	28	57.1	388	2	F70133	flagellar-associa
454	28	57.1	266	2	T35212	probable RNA polym	527	28	57.1	391	1	LPRTA4	apolipoprotein A-I
455	28	57.1	267	2	A69319	thiamin biosynthes	528	28	57.1	394	2	B75404	probable amidase -
456	28	57.1	268	2	T30630	hypothetical prote	529	28	57.1	396	2	T11709	nifs protein homol
457	28	57.1	272	1	SUGQBG	gliadin beta chain	530	28	57.1	397	2	S46691	exopolysphatase
458	28	57.1	273	2	E87430	flagellin FljL lim	531	28	57.1	403	2	AC1960	two-component resp
459	28	57.1	274	2	T16003	hypothetical prote	532	28	57.1	411	2	AB0199	peptidase T (EC 3.
460	28	57.1	282	2	G84081	hypothetical prote	533	28	57.1	412	2	F97911	probable UDPglucos
461	28	57.1	285	2	D70892	hypothetical prote	534	28	57.1	419	2	E70489	processing protein
462	28	57.1	285	2	S51247	thioredoxin homolo	535	28	57.1	420	2	A83948	comptence-damage
463	28	57.1	291	2	A11724	ABC transporter (A	536	28	57.1	426	2	C85755	probable oxidoredu
464	28	57.1	293	2	AC2480	hypothetical prote	537	28	57.1	426	2	F90863	probable oxidoredu
465	28	57.1	297	2	D69327	succinoglycan bios	538	28	57.1	426	2	H64878	probable oxidoredu
466	28	57.1	301	1	RGECGT	GTP-binding protei	539	28	57.1	430	2	AC2707	two component resp
467	28	57.1	301	2	E95977	UTP-glucose-1-phos	540	28	57.1	433	2	AI0882	probable oxidoredu

541	28	57.1	433	2	S25194	zuotin - yeast (sa	614	28	57.1	765	2	T15447	hypothetical prote
542	28	57.1	435	2	A2582	heat shock chapero	615	28	57.1	777	2	F84985	endopeptidase la (
543	28	57.1	435	2	D57364	heU protein (AE00	616	28	57.1	783	2	AE1961	Zam protein limpor
544	28	57.1	435	2	F96939	TPR repeats contai	617	28	57.1	786	2	A69308	DNA gyrase, subuni
545	28	57.1	435	2	A11017	probable type-I se	618	28	57.1	787	2	H86707	formate C-acetyltr
546	28	57.1	435	2	B87592	hypothetical prote	619	28	57.1	787	2	A71642	penicillin-binding
547	28	57.1	435	2	T30114	hypothetical prote	620	28	57.1	789	2	S53397	actin-interacting
548	28	57.1	436	2	C83449	probable oxidoredu	621	28	57.1	804	2	A39972	segment S5 protein
549	28	57.1	437	2	T23652	alpha-1,3-mannosyl	622	28	57.1	808	2	S64492	hypothetical prote
550	28	57.1	437	2	T02689	hypothetical prote	623	28	57.1	832	2	S56230	hypothetical prote
551	28	57.1	439	2	T14872	ABC exporter outer	624	28	57.1	842	2	T16198	hypothetical prote
552	28	57.1	441	1	B33862	transcription regu	625	28	57.1	842	2	C28667	DNA mismatch repai
553	28	57.1	441	1	B33862	cysteine synthase	626	28	57.1	852	2	F44020	hypothetical prote
554	28	57.1	442	2	T47800	response regulator	627	28	57.1	857	2	E98107	DNA mismatch repai
555	28	57.1	445	2	C97489	triacylglycerol li	628	28	57.1	865	2	A83946	DNA mismatch repai
556	28	57.1	449	1	L1PG	hypothetical prote	629	28	57.1	872	2	B89904	ATP-dependent Clp
557	28	57.1	453	2	S2690	membrane-bound lyl	630	28	57.1	874	2	AC3070	hypothetical prote
558	28	57.1	455	2	A80534	hypothetical prote	631	28	57.1	879	2	T19919	probable polyprote
559	28	57.1	457	2	H85013	hypothetical prote	632	28	57.1	884	2	E96452	endopeptidase clp
560	28	57.1	462	2	A64416	adenylosuccinate 1	633	28	57.1	887	2	F98216	hypothetical prote
561	28	57.1	462	2	T01549	hypothetical prote	634	28	57.1	909	2	AC2365	hypothetical prote
562	28	57.1	465	2	C43357	triacylglycerol li	635	28	57.1	917	2	S64100	probable membrane
563	28	57.1	466	1	P2WLRI	L2 protein - rhesu	636	28	57.1	919	2	B72263	isoleucine-cRNA li
564	28	57.1	467	2	E70976	hypothetical prote	637	28	57.1	925	2	T29311	hypothetical prote
565	28	57.1	469	1	P2WL35	L2 protein - human	638	28	57.1	931	2	T31731	probable preprotei
566	28	57.1	469	1	S36525	membrane-bound lyl	639	28	57.1	940	2	D70006	probable secA prot
567	28	57.1	475	2	A80132	splicing factor U2	640	28	57.1	949	2	B70592	hypothetical prote
568	28	57.1	475	2	S20250	hypothetical prote	641	28	57.1	955	2	T33040	hypothetical prote
569	28	57.1	480	2	D75050	RUB3 protein - yea	642	28	57.1	956	2	A56920	glutactin precurs
570	28	57.1	484	2	S60943	alkaline exonuclea	643	28	57.1	968	2	T01733	hypothetical prote
571	28	57.1	485	2	T03131	splicing factor U2	644	28	57.1	968	2	G90636	probable ATP-depen
572	28	57.1	492	2	S22646	hypothetical prote	645	28	57.1	968	2	G85487	probable ATP-depen
573	28	57.1	495	2	H71933	H+-transporting tw	646	28	57.1	975	1	C64727	kinesin heavy chai
574	28	57.1	504	2	S73163	H+-transporting tw	647	28	57.1	975	1	A31497	regulatory protein
575	28	57.1	505	2	S39520	alkyl hydroperoxid	648	28	57.1	978	1	RGBY13	probable pre-tRNA
576	28	57.1	509	2	B86667	competence protein	649	28	57.1	978	1	T40803	cation efflux syst
577	28	57.1	509	2	A12316	hypothetical prote	650	28	57.1	1020	2	B82604	hypothetical prote
578	28	57.1	509	2	T06300	conserved cytohet	651	28	57.1	1044	2	T00342	carbamoylphosphate
579	28	57.1	511	2	AG0261	probable hypokinin	652	28	57.1	1073	2	B83051	hypothetical prote
580	28	57.1	515	2	E84577	argininosuccinate	653	28	57.1	1075	2	T34223	hypothetical prote
581	28	57.1	517	2	T50801	glucose-6-phosphat	654	28	57.1	1091	2	F83928	hypothetical prote
582	28	57.1	523	1	DEFFG6	glucose-6-phosphat	655	28	57.1	1155	2	G84332	FUN12/BIF-2 family
583	28	57.1	524	1	A31854	lanosterol 14alpha	656	28	57.1	1221	2	A44978	DNA topoisomerase
584	28	57.1	528	1	T04722	hypothetical prote	657	28	57.1	1252	2	B42771	reticulocyte-bind
585	28	57.1	531	2	D86432	hypothetical prote	658	28	57.1	1255	2	B97104	hypothetical prote
586	28	57.1	540	2	G69526	methyalmalonyl-CoA	659	28	57.1	1296	2	I40645	botulinum neurotox
587	28	57.1	548	1	C75202	dipeptide abc tran	660	28	57.1	1418	2	T15232	hypothetical prote
588	28	57.1	556	2	T28949	hypothetical prote	661	28	57.1	1436	2	S67655	probable membrane
589	28	57.1	564	2	A87384	cell fusion glycop	662	28	57.1	1495	2	S27001	alpha-1-macroglobu
590	28	57.1	574	1	VGNZBS	glycyl-tRNA synth	663	28	57.1	1500	2	A42210	alpha-1-macroglobu
591	28	57.1	574	2	A99189	hypothetical prote	664	28	57.1	1553	2	T03301	rab3 effector prot
592	28	57.1	581	2	T29830	hypothetical prote	665	28	57.1	1657	2	T19536	hypothetical prote
593	28	57.1	587	2	A31776	lactose permease -	666	28	57.1	1787	2	G97222	hypothetical prote
594	28	57.1	592	2	T52139	LRR-containing F-b	667	28	57.1	1846	2	T33079	hypothetical prote
595	28	57.1	602	2	T50974	related to hxB pro	668	28	57.1	2150	2	T08165	RNAI polypeptide -
596	28	57.1	603	1	QQBED1	probable hemolysin	669	28	57.1	2241	2	T20971	hypothetical prote
597	28	57.1	611	2	H82351	deoxyxylulose-5-ph	670	28	57.1	2261	2	T20978	spectrin beta chai
598	28	57.1	619	2	A97156	ABC transporter, A	671	28	57.1	2291	1	A46147	hypothetical prote
599	28	57.1	619	2	F82391	ABC transporter AT	672	28	57.1	2291	1	A46147	hypothetical prote
600	28	57.1	626	2	A13310	hypothetical prote	673	28	57.1	3068	1	A44062	genome polyprotein
601	28	57.1	639	2	T06735	hypothetical prote	674	28	57.1	5107	2	T29144	hypothetical prote
602	28	57.1	645	2	F96630	hypothetical prote	675	27	56.1	433	2	T21989	hypothetical prote
603	28	57.1	653	2	E84206	methyalmalonyl-CoA	676	27	55.1	36	2	A83870	complement C5a - r
604	28	57.1	657	2	G71602	protein with DnaJ	677	27	55.1	77	2	A57689	hypothetical prote
605	28	57.1	668	2	C71496	probable exinuclea	678	27	55.1	78	2	B5193	hypothetical prote
606	28	57.1	668	2	C96814	hypothetical prote	679	27	55.1	78	2	B90999	probable phosphotr
607	28	57.1	676	2	D81654	exinuclease ABC c	680	27	55.1	93	2	G81172	probable phosphotr
608	28	57.1	689	2	E88998	hypothetical prote	681	27	55.1	93	2	G86018	probable Pts syste
609	28	57.1	701	1	B44259	kinesin-related pr	682	27	55.1	95	2	AG0964	hypothetical prote
610	28	57.1	701	2	T20892	hypothetical prote	683	27	55.1	101	2	H72542	hypothetical prote
611	28	57.1	707	2	T29559	hypothetical prote	684	27	55.1	102	2	E97788	hypothetical prote
612	28	57.1	754	2	H84710	Mutator-like trans	685	27	55.1	105	1	CCMST	cytochrome c, test
613	28	57.1	760	2	T28224	ORF MSV063 probabl	686	27	55.1	107	2	T42918	hypothetical prote

687 27 55.1 107 2 T42903 hypothetical prote
688 27 55.1 108 2 S61627 hypothetical prote
689 27 55.1 108 2 T4634 hypothetical prote
690 27 55.1 112 2 D90078 conserved hypotet
691 27 55.1 113 2 C71223 hypothetical prote
692 27 55.1 115 2 C64578 conserved hypotet
693 27 55.1 118 2 B64630 hypothetical prote
694 27 55.1 121 2 C96932 transcription regu
695 27 55.1 123 2 C81851 ribosomal-binding
696 27 55.1 123 2 B81091 ribosome-binding f
697 27 55.1 124 2 C90423 hypothetical prote
698 27 55.1 126 2 B86778 regulatory protein
699 27 55.1 129 2 A81725 transcription regu
700 27 55.1 131 2 A11354 transcription regu
701 27 55.1 133 2 A75314 ribosomal protein
702 27 55.1 133 2 F90037 hypothetical prote
703 27 55.1 133 2 AC0427 conserved hypotet
704 27 55.1 134 1 J01889 AL3 protein - toma
705 27 55.1 135 2 G64395 hypothetical prote
706 27 55.1 137 2 G95102 hypothetical prote
707 27 55.1 138 1 C64005 hypothetical prote
708 27 55.1 140 2 A48552 protein US420 - in
709 27 55.1 151 2 H69078 ferripyochelin bin
710 27 55.1 151 2 T01176 hypothetical prote
711 27 55.1 154 2 T17580 hypothetical prote
712 27 55.1 157 2 A22434 hypothetical prote
713 27 55.1 158 1 S38990 glycine reductase
714 27 55.1 158 2 G70031 mutator Mutr prote
715 27 55.1 166 2 C71874 hypothetical prote
716 27 55.1 175 2 E99992 hypothetical prote
717 27 55.1 175 2 A81657 hypothetical prote
718 27 55.1 175 2 A81285 hypothetical prote
719 27 55.1 177 2 T16280 hypothetical prote
720 27 55.1 177 2 AG1831 hypothetical prote
721 27 55.1 178 2 H70345 conserved hypotet
722 27 55.1 181 2 A70178 conserved hypotet
723 27 55.1 182 2 I67436 interleukin-1-beta
724 27 55.1 183 1 NKVLC3 core antigen - hep
725 27 55.1 185 2 C87646 transcription anti
726 27 55.1 186 1 Q3EYR hypothetical 20.6K
727 27 55.1 186 2 H91128 hypothetical prote
728 27 55.1 186 2 G85973 hypothetical prote
729 27 55.1 186 2 B69389 conserved hypotet
730 27 55.1 188 2 B82584 transcription regu
731 27 55.1 188 2 A11225 hypothetical prote
732 27 55.1 188 2 F82956 conserved hypotet
733 27 55.1 189 1 IVB011 interferon alpha-I
734 27 55.1 189 1 IVB01C interferon alpha-I
735 27 55.1 189 1 IVB01D interferon alpha-I
736 27 55.1 190 2 E90143 conserved hypotet
737 27 55.1 192 2 S12314 hypothetical prote
738 27 55.1 193 2 D70046 conserved hypotet
739 27 55.1 193 2 S10911 hypothetical prote
740 27 55.1 199 2 T19584 hypothetical prote
741 27 55.1 200 2 T32485 hypothetical prote
742 27 55.1 203 2 D86660 30S ribosomal prot
743 27 55.1 204 2 T10766 patatin-like latex
744 27 55.1 206 2 A81588 hypothetical prote
745 27 55.1 206 2 AC1651 hypothetical prote
746 27 55.1 206 2 AC3478 protein ybIS precu
747 27 55.1 207 2 T37464 probable glutathio
748 27 55.1 212 2 T31759 hypothetical prote
749 27 55.1 214 1 K1ECA adenylate kinase
750 27 55.1 214 2 AC0563 adenylate kinase
751 27 55.1 214 2 G90694 adenylate kinase
752 27 55.1 214 2 C85545 adenylate kinase
753 27 55.1 219 2 C70659 probable lipoprote
754 27 55.1 220 2 D70659 probable lipoprote
755 27 55.1 221 2 B70378 transcription regu
756 27 55.1 221 2 C64040 hypothetical prote
757 27 55.1 224 2 F69601 cytidylate kinase
758 27 55.1 226 2 AP2362 transcription regu
759 27 55.1 227 2 F72334 conserved hypotet

760 27 55.1 234 2 E86622 hypothetical prote
761 27 55.1 234 2 E72001 hypothetical prote
762 27 55.1 234 2 A13170 I55 family transpo
763 27 55.1 236 2 F97736 hypothetical prote
764 27 55.1 236 2 D71733 phosphoribosylamin
765 27 55.1 240 1 A70031 conserved hypotet
766 27 55.1 240 2 H75014 3-oxoacyl- (acyl-ca
767 27 55.1 241 2 A75065 hypothetical prote
768 27 55.1 242 2 A97117 tRNA- (guanine-N1) -
769 27 55.1 243 2 F81536 hypothetical prote
770 27 55.1 244 2 H69844 diadenosine tetrap
771 27 55.1 244 2 S73750 ATP-binding protei
772 27 55.1 252 2 AC1239 undecaprenyl dipho
773 27 55.1 252 2 AG1601 undecaprenyl dipho
774 27 55.1 252 2 D75193 hypothetical prote
775 27 55.1 253 2 beta-neoendorphin
776 27 55.1 254 1 DFHU proteasome endopep
777 27 55.1 255 1 DFFG beta-neoendorphin
778 27 55.1 257 2 H70198 thymidylate kinase
779 27 55.1 258 2 AF2287 group 3 sigma 37-t
780 27 55.1 258 2 JC6318 beta-neoendorphin
781 27 55.1 259 1 CTXLPB corticotropin / li
782 27 55.1 260 1 B42882 motility protein (
783 27 55.1 261 2 G70195 pyridoxal kinase
784 27 55.1 264 2 G70195 indole-3-glycerol-
785 27 55.1 266 1 F64414 extragenic suppress
786 27 55.1 266 2 AH3017 shikimate 5-dehydr
787 27 55.1 266 2 H98266 hypothetical prote
788 27 55.1 269 2 D90173 UTP-glucose-1-phos
789 27 55.1 269 2 T18145 flagellin fljO fim
790 27 55.1 273 2 G87347 hypothetical prote
791 27 55.1 273 2 T48436 purine nucleoside
792 27 55.1 275 2 cell division inhi
793 27 55.1 277 2 C70351 ABC-type sulfate t
794 27 55.1 278 2 H70404 flagellar synthesi
795 27 55.1 279 2 A96913 conserved hypotet
796 27 55.1 280 2 B83463 conserved hypotet
797 27 55.1 281 2 AH1738 acetyl-coenzyme A
798 27 55.1 289 2 E71875 conserved hypotet
799 27 55.1 289 2 F64638 conserved hypotet
800 27 55.1 291 2 AB1384 conserved hypotet
801 27 55.1 291 2 AD1759 conserved hypotet
802 27 55.1 291 2 AD1303 conserved hypotet
803 27 55.1 291 2 AD1675 phase-related prot
804 27 55.1 292 2 G84092 adenocortical est
805 27 55.1 296 2 A40111 homoserine kinase
806 27 55.1 297 2 A70413 hypothetical prote
807 27 55.1 301 2 E96630 hypothetical prote
808 27 55.1 301 2 T51435 hypothetical prote
809 27 55.1 301 2 A70335 hypothetical prote
810 27 55.1 308 2 E84374 prephenate dehydro
811 27 55.1 309 2 A13000 tyrc protein (mpo
812 27 55.1 310 2 G98282 hypothetical prote
813 27 55.1 311 2 H98278 hypothetical prote
814 27 55.1 311 2 A13004 hypothetical prote
815 27 55.1 312 2 I40383 malate dehydrogena
816 27 55.1 313 2 C96926 endoglucanase, ami
817 27 55.1 313 2 H64144 hypothetical prote
818 27 55.1 314 2 S00937 repE protein - Sta
819 27 55.1 314 2 AG3108 conserved hypotet
820 27 55.1 316 2 T08807 mitochondrial proc
821 27 55.1 317 2 T25560 hypothetical prote
822 27 55.1 318 2 E72676 hypothetical prote
823 27 55.1 319 2 T41795 ACNPPV orf51 - Bom
824 27 55.1 319 2 A47499 thiamine diphospho
825 27 55.1 321 2 F82296 conserved hypotet
826 27 55.1 321 2 S73340 hypothetical prote
827 27 55.1 324 1 S73340 protein gp49 [Bact
828 27 55.1 324 2 E81364 hypothetical prote
829 27 55.1 326 2 E98178 deacetylase, proba
830 27 55.1 327 2 G90139 G protein pathway
831 27 55.1 330 2 A70422 hypothetical prote
832 27 55.1 331 2 A69113 isocitrate dehydro

833	27	55.1	331	2	A12972	906	27	55.1	438	2	E55578	hypothetical prote
834	27	55.1	331	2	B98310	907	27	55.1	441	2	F86258	protein F5011.14 [
835	27	55.1	332	2	C64073	908	27	55.1	444	2	B81825	glutamate dehydrog
836	27	55.1	332	2	T28882	909	27	55.1	444	2	H81050	glutamate dehydrog
837	27	55.1	337	2	C90400	910	27	55.1	444	2	B69530	adenylosuccinate l
838	27	55.1	338	2	H72679	911	27	55.1	446	2	D82730	adenosylhomocyste
839	27	55.1	338	2	T23365	912	27	55.1	446	2	A83213	probable ATP-depen
840	27	55.1	341	2	D64159	913	27	55.1	448	2	G89933	hypothetical prote
841	27	55.1	342	2	D95906	914	27	55.1	449	2	G81419	hypothetical prote
842	27	55.1	342	2	T43410	915	27	55.1	453	1	FGHUGB	fibrinogen gamma-B
843	27	55.1	342	2	S42885	916	27	55.1	458	2	G83690	hypothetical prote
844	27	55.1	342	2	S23438	917	27	55.1	460	2	F83953	hypothetical prote
845	27	55.1	342	2	T23224	918	27	55.1	465	2	C75028	h1-transporing AT
846	27	55.1	345	2	E70484	919	27	55.1	465	2	F71213	probable H(+)-tran
847	27	55.1	352	2	S38147	921	27	55.1	467	2	B72331	purine-binding che
848	27	55.1	352	2	D72264	922	27	55.1	467	2	P2WL33	L2 protein - human
849	27	55.1	352	2	C98155	923	27	55.1	473	1	F2WLHS	hypothetical prote
850	27	55.1	353	2	E70100	924	27	55.1	480	2	C69853	probable type I se
851	27	55.1	353	2	A81737	925	27	55.1	483	2	A85723	minor capsid prote
852	27	55.1	357	2	A13132	926	27	55.1	483	2	D64906	probable tagaturon
853	27	55.1	360	2	E69086	927	27	55.1	483	2	H90894	altronate oxidore
854	27	55.1	365	2	B54128	928	27	55.1	484	2	AD0699	altronate oxidore
855	27	55.1	366	2	G71179	929	27	55.1	486	2	T37572	probable metallopr
856	27	55.1	367	2	T06131	930	27	55.1	491	2	T23524	hypothetical prote
857	27	55.1	370	2	D83793	931	27	55.1	492	2	F70326	conserved hypoteth
858	27	55.1	371	2	B85090	932	27	55.1	493	2	G84594	probable diacygly
859	27	55.1	375	2	I39169	933	27	55.1	493	2	G54382	acetylactate synth
860	27	55.1	375	2	I48686	934	27	55.1	494	2	T28660	probable adhesin p
861	27	55.1	375	2	A46727	935	27	55.1	499	2	E84776	hypothetical prote
862	27	55.1	375	2	S13025	936	27	55.1	503	2	S78321	H1-transporing tw
863	27	55.1	378	2	T00481	937	27	55.1	506	2	S71591	aspartic proteinas
864	27	55.1	382	1	WZV216	938	27	55.1	511	2	S58399	cellulagin I synt
865	27	55.1	382	2	B72158	939	27	55.1	513	2	G83725	GMP synthetase gua
866	27	55.1	382	2	S33075	940	27	55.1	520	1	D44771	3-oxoacid CoA-tran
867	27	55.1	382	2	T28498	941	27	55.1	521	2	D64434	lanosterol 14alpha
868	27	55.1	383	2	T50502	942	27	55.1	528	1	O4CK51	probable acid-CoA
869	27	55.1	383	2	E81156	943	27	55.1	529	2	G72605	probable DNA repai
870	27	55.1	383	2	E81950	944	27	55.1	529	2	H71450	probable ribose/ga
871	27	55.1	386	2	C90190	945	27	55.1	533	1	D71338	hypothetical prote
872	27	55.1	392	2	S64914	946	27	55.1	533	2	E86412	aldehyde dehydroge
873	27	55.1	397	2	A72408	947	27	55.1	538	2	T06683	hypothetical prote
874	27	55.1	398	2	G71733	948	27	55.1	539	2	A39481	hypothetical prote
875	27	55.1	399	2	T11954	949	27	55.1	540	2	T21647	hypothetical prote
876	27	55.1	400	2	D72277	950	27	55.1	541	2	S51799	nucleoporin NUP57
877	27	55.1	401	2	T19258	951	27	55.1	546	2	A39600	conserved hypoteth
878	27	55.1	402	2	C95035	952	27	55.1	548	2	B81698	pyrophosphate-fruc
879	27	55.1	403	2	T39846	953	27	55.1	548	2	G82286	phosphate ABC tran
880	27	55.1	404	2	F82092	954	27	55.1	549	2	F64640	conserved hypoteth
881	27	55.1	405	2	E90154	955	27	55.1	553	2	B64622	osmoprotection pro
882	27	55.1	406	2	F70019	956	27	55.1	555	2	D95377	probable cerebrosi
883	27	55.1	411	2	I39809	957	27	55.1	556	2	T23172	hypothetical prote
884	27	55.1	413	2	F70211	958	27	55.1	557	2	T45135	chaperone protein
885	27	55.1	416	2	A48249	959	27	55.1	565	2	T29718	hypothetical prote
886	27	55.1	419	2	A27539	960	27	55.1	565	2	JF0160	sodium bicarbonate
887	27	55.1	420	2	S76601	961	27	55.1	567	1	D43719	urease (EC 3.5.1.5
888	27	55.1	420	2	C69184	962	27	55.1	567	1	S08480	urease (EC 3.5.1.5
889	27	55.1	420	2	AF2464	963	27	55.1	570	2	T38148	phosphatidyl synth
890	27	55.1	421	2	H89916	964	27	55.1	572	2	S68119	lactase (EC 1.10.3
891	27	55.1	421	2	AE2473	965	27	55.1	575	2	S44018	lactase (EC 1.10.3
892	27	55.1	423	2	D81248	966	27	55.1	576	2	T38666	probable trp-asp r
893	27	55.1	423	2	AC3553	967	27	55.1	576	2	T38666	hypothetical prote
894	27	55.1	425	2	T46698	968	27	55.1	581	2	H70125	translation elonga
895	27	55.1	427	1	A39744	969	27	55.1	582	2	B70389	transport ATP-bind
896	27	55.1	430	2	JQ1392	970	27	55.1	582	2	D82146	probable protein-e
897	27	55.1	430	2	T37198	971	27	55.1	583	1	H71326	acetylactate synth
898	27	55.1	431	2	JC7730	972	27	55.1	584	2	B72362	malate oxidoreduct
899	27	55.1	432	2	T33118	973	27	55.1	588	2	A82232	hypothetical prote
900	27	55.1	434	2	F69177	974	27	55.1	592	2	T34446	hypothetical prote
901	27	55.1	434	2	T43197	975	27	55.1	596	2	F82174	methyl-accepting c
902	27	55.1	435	2	B72418	976	27	55.1	598	2	T01720	hypothetical prote
903	27	55.1	437	1	FGHUG	977	27	55.1	600	2	C97790	hypothetical prote
904	27	55.1	437	2	AH1271	978	27	55.1				
905	27	55.1	437	2	AB1634	979	27	55.1				

979 27 55.1 601 2 AF1192 heat shock protein
980 27 55.1 611 1 S62811 oligonucleotide
981 27 55.1 611 1 S62811 glucose inhibited
982 27 55.1 622 2 F97054 Fe-S oxidoreductase
983 27 55.1 623 2 T07933 polyadenylate-binding
984 27 55.1 632 2 B69310 mRNA 3'-end process
985 27 55.1 632 2 S58152 hypothetical prote
986 27 55.1 639 2 H97807 exonuclease ABC s
987 27 55.1 640 2 T29784 hypothetical prote
988 27 55.1 644 2 S51765 sulfate transport
989 27 55.1 647 2 F95189 threonyl-tRNA synt
990 27 55.1 657 2 A11882 nitrate transport
991 27 55.1 658 2 A33598 endoglucanase 3 (E
992 27 55.1 659 2 C85016 hypothetical prote
993 27 55.1 660 2 G98055 threonine-tRNA lig
994 27 55.1 663 2 F82160 conserved hypothet
995 27 55.1 663 2 A33618 heat shock protein
996 27 55.1 669 2 D72278 endo-1,4-beta-mann
997 27 55.1 673 2 T06294 hypothetical prote
998 27 55.1 677 2 G59895 formate dehydrogen
999 27 55.1 682 2 B84021 two-component sens
1000 27 55.1 686 2 D71292 probable DNA recom

ALIGNMENTS

RESULT 1
G71827
hypothetical protein jhp1271 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: G71827
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Milla, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1164 <ARN>
A:Cross-references: UNIPROT:Q9ZJN3; UNIPARC:UPI00000D72F4; GB:AE001550; GB:AE001439; NID
A:Experimental source: strain J99

Query Match 81.6%; Score 40; DB 2; Length 1164;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 740 ESLKEQVLKY 749

RESULT 2
S73600
probable GTP-binding protein spg - Mycoplasma pneumoniae (strain ATCC 29342)
N:Alternate names: hypothetical protein D02_orf291
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C>Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 31-Dec-2004
C:Accession: S73600
R:Hammelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-291 <HM>
A:Cross-references: UNIPROT:P75210; UNIPARC:UPI0000012A0F3; EMBL:AE000027; GB:U00089; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:
A:Gene: spg
A:Genetic code: SGC3
C:Superfamily: GTP-binding protein era homolog, bacteria type; translation elongation fa
C:Keywords: nucleotide binding; P-loop
F:10-17/Region: nucleotide-binding motif A (P-loop)

Query Match 77.6%; Score 38; DB 2; Length 291;
Best Local Similarity 60.0%; Pred. No. 9.2;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 191 EALREQLIKY 200

RESULT 3
A47305
translation initiation factor eIF-5 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A47305
R:Das, K.; Chevesich, J.; Maitra, U.
Proc. Natl. Acad. Sci. U.S.A. 90, 3058-3062, 1993
A:Title: Molecular cloning and expression of cDNA for mammalian translation initiation f
A:Reference number: A47305; MUID:93219424; PMID:8464924
A:Accession: A47305
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-429 <DAS>
A:Cross-references: UNIPROT:Q07205; UNIPARC:UPI000012D3A8; GB:L11651; NID:g294544; PIDN:
A:Experimental source: insulinoma
A:Note: sequence extracted from NCBI backbone (NCBIN:128800, NCBI:P128802)

Query Match 75.5%; Score 37; DB 2; Length 429;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 285 EKIREQIKKY 294

RESULT 4
S54048
ribosomal protein S13.e, cytosolic - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D4252; protein YBRC151i; protein YD9609.18; protein YDR064w;
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S54048; S58844; S61753; S11578; S67880
R:Hunt, S.; Bowman, S.; Harris, D.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54031
A:Accession: S54048
A:Molecule type: DNA
A:Residues: 1-151 <HUN>
A:Cross-references: UNIPROT:P05756; UNIPARC:UPI000004F8EC; EMBL:Z49209; NID:g798897; PIDN:
R:Brandt, P.; Otto, B.; Ramlow, S.; Blocker, H.
submitted to the EMBL Data Library, January 1995
A:Reference number: S58832
A:Accession: S58844
A:Molecule type: DNA
A:Residues: 1-151 <BRA>
A:Cross-references: UNIPARC:UPI000004F8EC; EMBL:X84162; NID:g706817; PIDN:CAAS8980.1; PII:
R:Brandt, P.; Ramlow, S.; Otto, B.; Blocker, H.
Yeast 12, 85-90, 1996
A:Title: Nucleotide sequence analysis of a 32,500 bp region of the right arm of Saccharo
A:Reference number: S61741; MUID:96381250; PMID:8789263
A:Accession: S61753
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-151 <BRW>
A:Cross-references: UNIPARC:UPI000004F8EC; EMBL:X84162; NID:g706817; PIDN:CAAS8980.1; PII:

A;Accession: T30254
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1234 <TAK>
A;Cross-references: UNIPROT:O62315; UNIPARC:UPI000000406B; EMBL:D31967; NID:g780143; PID:
A;Experimental source: strain 129/Ola
C;Genetics:
A;Gene: jumonji
C;Function:
A;Description: may be essential for normal morphogenesis of the neural tube

Query Match 73.5%; Score 36; DB 2; Length 1234;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|:
Db 573 ESVRAQVEKY 582

RESULT 7
A29924
acetyl-CoA carboxylase (EC 6.4.1.2), hepatic - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A29924; A29337; A27903
R;Takai, T.; Yokoyama, C.; Wada, K.; Tanabe, T.
J. Biol. Chem. 263, 2651-2657, 1988
A;Title: Primary structure of chicken liver acetyl-CoA carboxylase deduced from cDNA seq
A;Reference number: A29924; MUID:88139305; PMID:2893793
A;Accession: A29924
A;Molecule type: mRNA
A;Residues: 1-2324 <TAK1>
A;Cross-references: UNIPROT:P11029; UNIPARC:UPI0000127CAB; GB:J03541; NID:g211567; PID:
R;Takai, T.; Wada, K.; Tanabe, T.
FEBS Lett. 212, 98-102, 1987
A;Title: Primary structure of the biotin-binding site of chicken liver acetyl-CoA carboxylase
A;Reference number: A91375; MUID:87106011; PMID:2879745
A;Accession: A29337
A;Molecule type: mRNA
A;Residues: 493-820 <TAK2>
A;Cross-references: UNIPARC:UPI0000171282; GB:X05019; NID:g63021; PID:CAA28675.1; PID:9
A;Accession: A27903
A;Molecule type: mRNA
A;Residues: 493-552,554-783,'RSPS',789-820 <TAK3>
A;Cross-references: UNIPARC:UPI000017312E
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase; liver
F;120-620/Domain: biotin carboxylase homology <BCH>
F;747-819/Domain: lipoyl/biotin-binding homology <LBP>
F;786/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 1; Length 2324;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|:
Db 1467 ESVRSVMVMRY 1476

RESULT 8
A35578
acetyl-CoA carboxylase (EC 6.4.1.2) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A35578; A37119; I59145; I70069; I70070; I55305
R;Lopez-Casillas, F.; Bai, D.H.; Luo, X.; Kong, I.S.; Hermanson, M.A.; Kim, K.H.
Proc. Natl. Acad. Sci. U.S.A. 85, 5784-5788, 1988
A;Title: Structure of the coding sequence and primary amino acid sequence of acetyl-CoA
A;Reference number: A35578; MUID:88320328; PMID:2901088
A;Accession: A35578
A;Molecule type: mRNA

R;Otaka, E.; Higo, K.; Osawa, S.
Biochemistry 21, 4545-4550, 1982
A;Title: Isolation of seventeen proteins and amino-terminal amino acid sequences of eight
A;Reference number: S11575; MUID:81048950; PMID:6814480
A;Accession: S11578
A;Molecule type: protein
A;Residues: 2-24,'G',26-31,'C',33-35,'X',37-41 <OTA>
A;Cross-references: UNIPARC:UPI00001772A5
R;Bloecker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67587
A;Accession: S67880
A;Molecule type: DNA
A;Residues: 1-151 <BLO>
A;Cross-references: UNIPARC:UPI000004F8EC; EMBL:Z74360; MIPS:YDR064w; NID:g1431516; PID:
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:RPS13C
A;Cross-references: SGD:S0002471; MIPS:YDR064w
A;Map position: 4R
A;Introns: 7/3
C;Superfamily: rat ribosomal protein S13; eubacterial ribosomal protein S15 homology
C;Keywords: protein biosynthesis; ribosome
F;82-148/Domain: eubacterial ribosomal protein S15 homology <BS15>

Query Match 73.5%; Score 36; DB 2; Length 151;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|:
Db 31 ESVIEQIVKY 40

RESULT 5
F64816
ybiA protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: F64816
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F64816
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-160 <BLAT>
A;Cross-references: UNIPROT:P30176; UNIPARC:UPI000013A364; GB:AE0000182; GB:U000096; NID:9
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybiA

Query Match 73.5%; Score 36; DB 2; Length 160;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMK 9
|||:|:|:
Db 90 ESVREQVMR 98

RESULT 6
T30254
jumonji protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30254
R;Takeuchi, T.; Yamazaki, Y.; Katoh-Fukui, Y.; Tsuchiya, R.; Kondo, S.; Motoyama, J.; Hi
Genes Dev. 9, 1211-1222, 1995
A;Title: Gene trap capture of a novel mouse gene, jumonji, required for neural tube form
A;Reference number: Z20792; MUID:95278734; PMID:7758946

A;Residues: 1-2345 <LOP>
A;Cross-references: UNIPROT:P11497; UNIPARC:UPI0000127C7B; GB:J03808; NID:g202644; PIDN:
R;Kong, I.S.; Lopez-Casillas, F.; Kim, K.H.
J. Biol. Chem. 265, 13695-13701, 1990
A;Title: Acetyl-CoA carboxylase mRNA species with or without inhibitory coding sequence
A;Reference number: A37119; MUID:90337981; PMID:1974251
A;Accession: A37119
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1167-1200 <KON>
A;Cross-references: UNIPARC:UPI000017312F; GB:M55315
A;Experimental source: clone lambdaDH121
R;Luo, X.
Proc. Natl. Acad. Sci. U.S.A. 86, 4042-4046, 1989
A;Title: Structural features of the acetyl-CoA carboxylase gene: Mechanism for the gene
A;Reference number: I59145; MUID:89264558; PMID:2566999
A;Accession: I59145
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-74 <RES>
A;Cross-references: UNIPARC:UPI00001708B7; GB:M26731; NID:g202641; PIDN:AAA40652.1; PID:
A;Experimental source: hepatic
R;Lopez-Casillas, F.; Kim, K.
J. Biol. Chem. 264, 7176-7184, 1989
A;Title: Heterogeneity at the 5' end of rat acetyl-coenzyme A carboxylase mRNA: Lipogeni
A;Reference number: I55305; MUID:89214151; PMID:2565337
A;Accession: I70069
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE2>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26196; NID:g202649; PIDN:AAA40655.1; PID:
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE3>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26197; NID:g202651; PIDN:AAA40656.1; PID:
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-33 <RE4>
A;Cross-references: UNIPARC:UPI000017084D; GB:M26195; NID:g202647; PIDN:AAA40654.1; PID:
A;Experimental source: hepatic
C;Comment: This enzyme catalyzes the carboxylation of acetyl CoA to malonyl CoA and is t
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase
F;119-619/Domain: biotin carboxylase homology <BCH>
F;746-818/Domain: lipoyl/biotin-binding homology <LPB>
F;785/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 1; Length 2345;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 1489 ESVRSVMVRY 1498

RESULT 9
I38928
acetyl-CoA carboxylase (RC 6.4.1.2), hepatic - human
N;Alternate names: acetyl-Coenzyme A carboxylase
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
C;Accession: I38928
R;Abu-Elheiga, L.; Jayakumar, A.; Baldini, A.; Chirala, S.S.; Wakil, S.J.
Proc. Natl. Acad. Sci. U.S.A. 92, 4011-4015, 1995
A;Title: Human acetyl-CoA carboxylase: characterization, molecular cloning, and evidence
A;Reference number: I38928; MUID:95249602; PMID:7732023
A;Accession: I38928
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2346 <RES>

A;Cross-references: UNIPROT:Q13085; UNIPARC:UPI0000127C6D; EMBL:U19822; NID:g849082; PID:
A;Experimental source: HepG2 cells
C;Genetics: ACC
A;Gene: ACC
A;Map position: 17q12
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase; liver
F;120-620/Domain: biotin carboxylase homology <BCH>
F;747-819/Domain: lipoyl/biotin-binding homology <LPB>
F;786/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 2; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 1490 ESVRSVMVRY 1499

RESULT 10
G72358
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: G72358
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: G72358
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-486 <ARN>
A;Cross-references: UNIPROT:Q9WZ51; UNIPARC:UPI00000C13E8; GB:AE001733; GB:AE000512; NID:
A;Experimental source: strain MSB8
C;Genetics: TM0582
C;Superfamily: hypothetical protein MJ0966

Query Match 71.4%; Score 35; DB 2; Length 486;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 439 ESIRSVVQY 448

RESULT 11
S45907
myb-related protein REB1 - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YBR049c; protein YBR0502
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: S45907; S49504; A35924; S55847
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S45907
A;Molecule type: DNA
A;Residues: 1-810 <ALJ>
A;Cross-references: UNIPROT:P21538; UNIPARC:UPI0000053008; EMBL:Z35918; NID:g536279; PID:
R;Aljinovic, G.
submitted to the EMBL Data Library, October 1994
A;Description: Sequence and analysis of 24 kb on chromosome II of Saccharomyces cerevisi
A;Reference number: S49503
A;Accession: S49504
A;Molecule type: DNA
A;Residues: 1-810 <AL2>

A:Cross-references: UNIPARC:UPI0000053008; EMBL:Z46260; NID:G559942; PIDN:CAA86391.1; PI
A:Experimental source: strain S288C
R:Ju, Q.; Morrow, B.E.; Warner, J.R.
Mol. Cell. Biol. 10, 5226-5234, 1990
A:Title: REB1, a yeast DNA-binding protein with many targets, is essential for cell growth
A:Reference number: A35924; MUID:90377212; PMID:2204808
A:Accession: A35924
A:Molecule type: DNA
A:Residues: 1-55, 'N', 57-591, 'RAVFKNNFFNKSKMTMLRSC', 619-626, 'E', 628-635, 'E', 637-810
A:Cross-references: UNIPARC:UPI0000168DD6; EMBL:M58728; NID:G172371; PIDN:AAA4963.1; PI
R:Aljinovic, G.; Pohl, T.M.
Yeast 11, 475-479, 1995
A:Title: Sequence and analysis of 24 kb on chromosome II of *Saccharomyces cerevisiae*.
A:Reference number: S55846; MUID:95321020; PMID:7597852
A:Accession: S55847
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-810 <ALM>
A:Cross-references: UNIPARC:UPI0000053008; EMBL:Z46260; NID:G559942; PIDN:CAA86391.1; PI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Gene: SGD:REB1; MIPS:YBR049C
A:Cross-references: SGD:S0000253; MIPS:YBR049C
A:Map position: 2R
C:Superfamily: myb-related protein REB1; myb DNA-binding repeat homology
C:Keywords: DNA binding; nucleus; transcription regulation
F:470-519/Domain: myb DNA-binding repeat homology <MYB>

Query Match 71.4%; Score 35; DB 1; Length 810;
Best Local Similarity 70.0%; Pred. No. 98;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||| | | |
DB 13 ESVREAVLKY 22

RESULT 12
E69681
peptide synthetase ppsE - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: E69681
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69681
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1279 <KUN>
A:Cross-references: UNIPROT:O31827; UNIPARC:UPI00000604A5; GB:Z99113; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: ppsE
C:Superfamily: peptide synthetase ppsE; acetate-CoA ligase homology; acyl carrier protei
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:522-960/Domain: acetate-CoA ligase homology <ACL>
F:978-1046/Domain: acyl carrier protein homology <ACP>
F:1010/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 71.4%; Score 35; DB 1; Length 1279;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 9
|:|:|:| |
DB 1180 EAVREKVMK 1189

RESULT 13
G82176
Probable outer membrane protein VCL622 [imported] - *Vibrio cholerae* (strain N16961 serog
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82176
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, C.M.
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82176
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <HEI>
A:Cross-references: UNIPROT:Q9KRL5; UNIPARC:UPI00000C3053; GB:AE004240; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCL622
A:Map position: 1

Query Match 69.4%; Score 34; DB 2; Length 212;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:| |
DB 147 EAVRDQLLY 156

RESULT 14
B90544
Hypothetical protein MYPV 2580 [imported] - *Mycoplasma pulmonis* (strain UAB CTIP)
C:Species: *Mycoplasma pulmonis*
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 05-Oct-2004
C:Accession: B90544
R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: B90544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <KUR>
A:Cross-references: UNIPROT:Q98QV3; UNIPARC:UPI00000D45A0; GB:AL445566; PID:G14089671; F
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 2580
A:Genetic code: SG3
C:Superfamily: bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofo

Query Match 69.4%; Score 34; DB 2; Length 272;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:| |
DB 236 ESIKEALKY 245

RESULT 15
AD1062
protein kinase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (strain C

C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: This species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD1062
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD1062
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-357 <PAR>
A;Cross-references: UNIPARC:UPI000005A993; GB:AL513382; PIDN:CAD06945.1; PID:G16505592; C;Genetics:
A;Gene: STV4823

Query Match 69.4%; Score 34; DB 2; Length 357;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SVREQVMKY 10
| | | | | | |
Db 254 SVRQIQIAY 262

RESULT 16
E90514
hypothetical protein MYPV 0210 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: *Mycoplasma pulmonis*
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: E90514
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*
A;Reference number: A99512; MUID:21267165; PMID:11333084
A;Accession: E90514
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-455 <KUR>
A;Cross-references: UNIPROT:Q98R18; UNIPARC:UPI00000D4547; GB:AL445566; PID:G14089434; E; A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPV 0210
A;Genetic code: SGC3
C;Superfamily: dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase homology
C;Keywords: redox-active disulfide
F;41-46/Diulfide bonds: redox-active #status predicted

Query Match 69.4%; Score 34; DB 2; Length 455;
Best Local Similarity 75.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VREQVMKY 10
| | | | | | |
Db 214 VREVLKY 221

RESULT 17
T17590
probable glutamine-fructose-6-phosphate transaminase (isomerizing) (EC 2.6.1.16) - Chlorella virus PBCV-1
C;Species: *Chlorella virus PBCV-1*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17590
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17590
A;Status: Preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-595 <GRA>

A;Cross-references: UNIPROT:Q84421; UNIPARC:UPI00000F39A6; EMBL:U42580; NID:G4028896; PID: C;Genetics:
A;Note: AL00R
C;Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)
C;Keywords: aminotransferase; isomerase
F;2-595/Product: glutamine-fructose-6-phosphate transaminase (isomerizing) #status predicted; F;2/Active site: Cys #status predicted

Query Match 69.4%; Score 34; DB 2; Length 595;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VREQVMKY 10
| | | | | | |
Db 378 VRESIMKY 385

RESULT 18
S71091
acetyl-CoA carboxylase (EC 6.4.1.2) beta - human (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S71091
R;Widmer, J.; Passi, K.S.; Schlichter, S.C.; Whesler, K.S.; Crute, B.E.; King, N.; Nut Biochem. J. 316, 915-922, 1996
A;Title: Identification of a second human acetyl-CoA carboxylase gene.
A;Reference number: S71091; MUID:96265061; PMID:8670171
A;Accession: S71091
A;Status: Preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-786 <WID>
A;Cross-references: UNIPROT:O00763; UNIPARC:UPI000016B177; EMBL:U34591; NID:G1399289; PID: C;Genetics:
A;Gene: ACC-beta
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin- C;Keywords: ligase

Query Match 69.4%; Score 34; DB 2; Length 786;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | | | |
Db 286 ESVRVMYRY 295

RESULT 19
S54465
YTA12 protein precursor, mitochondrial - Yeast (*Saccharomyces cerevisiae*)
N;Alternate names: protein YMS82.14c; protein YMR089c; RCA1 protein
C;Species: *Saccharomyces cerevisiae*
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S54465; S46609; A55358; S48539
R;Gentles, S.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54451
A;Accession: S54465
A;Molecule type: DNA
A;Residues: 1-825 <GEN>
A;Cross-references: UNIPROT:P40341; UNIPARC:UPI0000133438; EMBL:Z49259; NID:G807956; PID: A;Experimental source: strain AB972
R;Schnall, R.; Mannhaupt, G.; Stucka, R.; Tauer, R.; Ehnlé, S.; Schwarzlose, C.; Vetter, Yeast 10, 1141-1155, 1994
A;Title: Identification of a set of yeast genes coding for a novel family of putative AT
A;Reference number: S46605; MUID:95274317; PMID:7754704
A;Accession: S46609
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-348, EL', 351-825 <SCH>
A;Cross-references: UNIPARC:UPI0000168FBI; EMBL:X81068; NID:G531753; PIDN:CAA56955.1; PID: R;Tagoloff, A.; Yue, J.; Jang, J.; Paul, M.F.
J. Biol. Chem. 269, 26144-26151, 1994
A;Title: A new member of a family of ATPases is essential for assembly of mitochondrial

A;Reference number: A55358; MUID:95014441; PMID:7929327
A;Accession: A55358
A;Molecule type: DNA
A;Residues: 1-652;V,654-825 <TZA>
A;Cross-references: UNIPARC:UPI00001689D8; GB:U09358; NID:G508232; PIDN:AAA62606.1; PID:
C;Genetics:
A;Gene: SGD:YTA12; RCA1
A;Cross-references: SGD:S0004695; MIPS:YMR089c
A;Map position: 13R
A;Genome: nuclear
C;Keywords: ATP; mitochondrion; nucleotide binding; P-loop; transmembrane protein
F;1-40/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;41-825/Product: YTA12 protein #status predicted <MAT>
F;178-194/Domain: transmembrane #status predicted <TM1>
F;299-311/Domain: transmembrane #status predicted <TM2>
F;362-574/Domain: FtsH/SSC18/CDC48-type ATP-binding domain homology <VATP>
F;388-395/Region: nucleotide-binding motif A (P-loop)

Query Match 69.4%; Score 34; DB 2; Length 825;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 71 EAIRQVEKY 80
|:|:|:|:|

RESULT 20
C38992
cadherin 4 precursor - human
N;Alternate names: R-cadherin
C;Species: Homo sapiens (man)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: C38992
R;Suzuki, S.; Sano, K.; Tanihara, H.
Cell Regul. 2, 261-270, 1991
A;Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous
A;Reference number: S24305; MUID:9128340; PMID:2059658
A;Accession: C38992
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-916 <SUZ>
A;Cross-references: UNIPROT:P55283; UNIPARC:UPI0000126D95; GB:L34059; NID:G506409; PIDN:
C;Genetics:
A;Gene: GDB:CDH4
A;Cross-references: GDB:622850
A;Map position: 16q24.1-16qter
C;Superfamily: cadherin; cadherin repeat homology
C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-166/Domain: propeptide #status predicted <PRO>
F;167-913/Product: R-cadherin #status predicted <MAT>
F;167-721/Domain: extracellular #status predicted <EXT>
F;169-274/Domain: cadherin repeat homology <CR1>
F;244-274/Region: cadherin binding #status predicted
F;277-389/Domain: cadherin repeat homology <CR2>
F;300-304/Domain: calcium binding #status predicted <CAB>
F;392-504/Domain: cadherin repeat homology <CR3>
F;507-612/Domain: cadherin repeat homology <CR4>
F;613-721/Domain: cadherin repeat homology <CR5>
F;722-753/Domain: transmembrane #status predicted <TMW>
F;754-913/Domain: intracellular #status predicted <INT>
F;870-885/Region: serine-rich
F;283,412,557,632,661,702/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 69.4%; Score 34; DB 2; Length 916;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 774 DDVREKILKY 783
|:|:|:|:|

RESULT 21

S41121
acetyl-CoA carboxylase (EC 6.4.1.2) - human
C;Species: Homo sapiens (man)
C;Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 11-Jan-2002
C;Accession: S41121
R;Ha, J.; Daniel, S.; Kong, I.S.; Park, C.K.; Tae, H.J.; Kim, K.H.
Eur. J. Biochem. 219, 297-306, 1994
A;Title: Cloning of human acetyl-CoA carboxylase cDNA.
A;Reference number: S41121; MUID:941139704; PMID:7905825
A;Accession: S41121
A;Molecule type: mRNA
A;Residues: 1-2339 <HAJ>
A;Cross-references: UNIPARC:UPI000011EDF5; EMBL:X68968; NID:G452315; PIDN:CAA48770.1; PI:
C;Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C;Keywords: biotin binding; ligase
F;120-620/Domain: biotin carboxylase homology <BCH>
F;747-819/Domain: lipoyl/biotin-binding homology <LPB>
F;786/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 69.4%; Score 34; DB 2; Length 2339;
Best Local Similarity 70.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 1490 ESVRYVMRY 1499
|:|:|:|:|

RESULT 22
B90408
hypothetical protein SSO2370 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: B90408
R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, J.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: B90408
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <KUR>
A;Cross-references: UNIPROT:Q97W67; UNIPARC:UPI0000064715; GB:AE006641; NID:G13815673; I:
C;Genetics:
A;Gene: SSO2370

Query Match 67.3%; Score 33; DB 2; Length 179;
Best Local Similarity 60.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 46 ESVRYTIMKY 55
|:|:|:|:|

RESULT 23
T25190
hypothetical protein T23G11.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25190
R;Gardner, A.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19993
A;Accession: T25190
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-243 <WIL>
A;Cross-references: UNIPROT:Q9XV10; UNIPARC:UPI0000179E6D; EMBL:Z81130; PIDN:CAB03414.1

A;Experimental source: clone T23G11
C;Genetics:
A;Gene: CRSP:T23G11.5
A;Map position: 1
A;Introns: 41/2; 82/3; 115/3; 199/2

Query Match 67.3%; Score 33; DB 2; Length 243;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 231 ESLSKVVMKY 240

RESULT 24
T03942
rRNA N-glycosidase (EC 3.2.2.22) - maize
N;Alternate names: ribosome-inactivating protein
C;Species: Zea mays (maize)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03942
R;Bass, H.W.; O'Brien, G.R.; Boston, R.S.
A;Title: Cloning and sequencing of a second ribosome-inactivating protein gene from maize
A;Reference number: Z15149; MUID:95241633; PMID:7724682
A;Accession: T03942
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-278 <BS>
A;Cross-references: UNIPROT:Q41851; UNIPARC:UPI000009D113; EMBL:L26305; NID:g496163; PID
C;Experimental source: strain W64A; leaf
C;Genetics:
A;Gene: RIP2
C;Function:
A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C;Keywords: glycosidase; hydrolase
F;27-277/Domain: rRNA N-glycosidase homology <RNG>

Query Match 67.3%; Score 33; DB 2; Length 278;
Best Local Similarity 66.7%; Pred. No. 82;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SVREQVMKY 10
|:|:|:|:|
Db 43 SVRKEVIKY 51

RESULT 25
A84341
hypothetical protein Vng1900c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C;Accession: A84341
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84341
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-282 <STO>
A;Cross-references: UNIPROT:Q9HMX7; UNIPARC:UPI00001320AD; GB:AE004437; NID:g10581342; F
C;Genetics:
A;Gene: VNG1900C
C;Superfamily: ATP-NAD(H) kinase

Query Match 67.3%; Score 33; DB 2; Length 282;
Best Local Similarity 60.0%; Pred. No. 83;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 103 EAVREVNRY 112

RESULT 26
G82581
GTP binding protein XF2247 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2004
C;Accession: G82581
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <SIM>
A;Cross-references: UNIPROT:Q9PB97; UNIPARC:UPI000012A10D; GB:AE004037; GB:AE003849; NID
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Reto, E.; Docena, C.; El-Dorri, H.; Pacinani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
; F.G.; Nunes, L.N.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2247
C;Superfamily: GTP-binding protein era homolog, bacteria type; translation elongation fa

Query Match 67.3%; Score 33; DB 2; Length 298;
Best Local Similarity 77.8%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 9
|:|:|:|:|
Db 197 ELVREQVMR 205

RESULT 27
H69486
conserved hypothetical protein AF1897 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69486
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, B.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69486
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-355 <KLE>
A;Cross-references: UNIPROT:O28382; UNIPARC:UPI0000056B64; GB:AE000972; GB:AE000782; NID

Query Match 67.3%; Score 33; DB 2; Length 355;
Best Local Similarity 70.0%; Pred. No. 1e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
||||| : ||
Db 43 ESVREIIEKY 52

RESULT 28
T43289
fructose-bisphosphate aldolase (EC 4.1.2.13) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43289
R:McDonagh, N.; Hayashi, Y.
Biochim. Biophys. Acta 1183, 550-552, 1994
A:Title: Molecular cloning and nucleotide sequencing of Schizosaccharomyces pombe homolo
A:Reference number: 222395; MUID:94114559; PMID:8286404
A:Accession: T43289
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-358 <MUT>
A:Cross-references: UNIPROT:P36580; UNIPARC:UPI0000169149; EMBL:DL17415; NID:G9398158; PID
C:Superfamily: fructose-bisphosphate aldolase II
C:Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis; pentose ph

Query Match 67.3%; Score 33; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
||||| : ||
Db 297 EGVRDYVLKY 306

RESULT 29
T39798
fructose-bisphosphate aldolase (EC 4.1.2.13) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39798
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21880
A:Accession: T39798
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-358 <MCD>
A:Cross-references: UNIPROT:P36580; UNIPARC:UPI0000125829; EMBL:AL109731; PIDN:CAB52034.
A:Experimental source: strain 972h-; cosmid c19C2
C:Genetics:
A:Gene: SPDB:SPBC19C2.07
A:Map position: 2
C:Superfamily: fructose-bisphosphate aldolase II
C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 67.3%; Score 33; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
||||| : ||
Db 297 EGVRDYVLKY 306

RESULT 30
AE2092
hypothetical protein all2292 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE2092
R:kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2092
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q8YUP3; UNIPARC:UPI000000CE3B3; GB:BA000019; PIDN:BA073991.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2292

Query Match 67.3%; Score 33; DB 2; Length 392;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
||||| : ||
Db 156 QSVREIILHY 165

RESULT 31
T12450
hypothetical protein DKFZp564N1916.1 - human
C:Species: Homo sapiens (man)
C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12450
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17523
A:Accession: T12450
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-431 <KOE>
A:Cross-references: UNIPROT:P55010; UNIPARC:UPI000016AC57; EMBL:AL080102
A:Experimental source: fetal brain; clone DKFZp564N1916
C:Genetics:
A:Note: DKFZp564N1916.1

Query Match 67.3%; Score 33; DB 2; Length 431;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
||||| : ||
Db 287 EKIREQIKEY 296

RESULT 32
F86841
iron-binding oxidase subunit [imported] - Lactococcus lactis subsp. lactis (strain IL14;
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86841
R:Polotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86841
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <STO>
A:Cross-references: UNIPROT:Q9CEV0; UNIPARC:UPI000006CABF; GB:AE005176; PID:g12724753;
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yrjC
C:Superfamily: conserved hypothetical iron-sulfur protein HP0138; ferredoxin 2 [4Fe-4S]

Query Match 67.3%; Score 33; DB 2; Length 490;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10

QY 1 ESVREQVMKY 9
|:|:|:|
Db 1010 EQIREIMKY 1018

RESULT 38
T23429
hypothetical protein K07G5.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23429
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19739
A:Accession: T23429
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1202 <WTL>
A:Cross-references: UNIPROT:Q211303; UNIPARC:UPI0000078B43; EMBL:Z71264; PIDN:CAA95830.1;
A:Experimental source: clone K07G5
C:Genetics:
A:Gene: CESP:K07G5.3
A:Map position: 1
A:Introns: 13/1; 104/3; 196/3; 260/2; 282/3; 307/2; 361/2; 410/2; 535/2; 603/2; 688/1; 7
C:Superfamily: Caenorhabditis elegans hypothetical protein K07G5.3

Query Match 67.3%; Score 33; DB 2; Length 1202;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 67 ESIKEHLMKF 76

RESULT 39
A30063
dihydropyridine receptor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A30063
R:Tanabe, T.; Takeshima, H.; Mikami, A.; Flockerzi, V.; Takahashi, H.; Kangawa, K.; Kojima
Nature 328, 313-318, 1987
A:Title: Primary structure of the receptor for calcium channel blockers from skeletal mu
A:Reference number: A30063; MUID:87258269; PMID:3037387
A:Accession: A30063
A:Molecule type: mRNA
A:Residues: 1-1873 <TAN>
A:Cross-references: UNIPROT:P07293; UNIPARC:UPI000005031B; GB:X05921; NID:g1547; PIDN:CA
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: phosphoprotein; transmembrane protein

Query Match 67.3%; Score 33; DB 2; Length 1873;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 825 ESVRNQILGY 834

RESULT 40
S69116
fibrinogen gamma chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Mar-1998 #sequence_revision 24-Apr-1998 #text_change 02-Jul-1998
C:Accession: S69116
R:Litvinovich, S.V.; Henschen, A.H.; Krieglstein, K.G.; Ingham, K.C.; Medved, L.V.
Eur. J. Biochem. 229, 605-614, 1995
A:Title: Structural and functional characterization of proteolytic fragments derived fro
A:Reference number: S69114; MUID:95278210; PMID:7758453
A:Accession: S69116

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-92 <LIT>
A:Cross-references: UNIPARC:UPI00001777E6
C:Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology

Query Match 65.3%; Score 32; DB 2; Length 92;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 26 KSMEEIMKY 35

RESULT 41

T27724
hypothetical protein ZK1251.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27724
R:McMurray, A.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z20411
A:Accession: T27724
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-138 <WTL>
A:Cross-references: UNIPROT:Q23434; UNIPARC:UPI00000793E7; EMBL:Z68222; PIDN:CAA92503.1;
A:Experimental source: clone ZK1251
C:Genetics:
A:Gene: CESP:ZK1251.7
A:Map position: 4
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK1251.7

Query Match 65.3%; Score 32; DB 2; Length 138;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 25 ENIROQATKY 34

RESULT 42

E97310
uncharacterized conserved domain seen in the bacterial Spot CAC3340 [imported] - Clostr
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: E97310
R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A36900; MUID:21359325; PMID:21359325
A:Accession: E97310
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <KUR>
A:Cross-references: UNIPROT:Q97DY0; UNIPARC:UPI00000D46F2; GB:AE001437; PIDN:AAK81272.1;
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3340
C:Superfamily: GTP pyrophosphokinase related protein

Query Match 65.3%; Score 32; DB 2; Length 217;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|:|:|:|
Db 20 EEMQEMIMKY 29

```
RESULT 43
H64321
hypothetical protein MJ0175 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: H64321
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reisch, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: H64321
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-222 <BUL>
A;Cross-references: UNIPROT:Q57639; UNIPARC:UPI00001393DD; GB:U67474; GB:L77117; NID:g15
C;Genetics:
A;Map position: REV178140-177472
A;Start codon: GTG

Query Match 65.3%; Score 32; DB 2; Length 222;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMK 9
:|||||
Db 32 DSVREILK 40

RESULT 44
S45880
SCO1 protein homolog SCO2 - Yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR024w; hypothetical protein YBR0308
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: S45880; S46555
R;Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45875
A;Accession: S45880
A;Molecule type: DNA
A;Residues: 1-301 <GR>
A;Cross-references: UNIPROT:P38072; UNIPARC:UPI000013566E; EMBL:Z35893; NID:9536231; PID
A;Experimental source: strain S288C
R;Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
Yeast 10(Suppl.A), S75-S80, 1994
A;Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A;Reference number: S46551; MUID:94378725; PMID:8091864
A;Accession: S46555
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <SM>
A;Cross-references: UNIPARC:UPI000013566E; EMBL:X76078; NID:G498748; PID:G498753
A;Experimental source: strain S288C
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Gene: SGD:SCO2
A;Cross-references: SGD:S0000228; MIPS:YBR024w
A;Map position: 2R
C;Keywords: mitochondrion; transmembrane protein

Query Match 65.3%; Score 32; DB 2; Length 301;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ESVREQVMK 10
:|||||
Db 273 EKIREIQAY 282
```

```
RESULT 45
S00909
replication initiation protein - Staphylococcus aureus plasmid pC223
C;Species: Staphylococcus aureus
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S00909
R;Ehret, M.; Matzura, H.
Nucleic Acids Res 16, 2045-2062, 1988
A;Title: Replication control of the Staphylococcus aureus chloramphenicol resistance pla
A;Reference number: S00909; MUID:88189801; PMID:3128771
A;Accession: S00909
A;Molecule type: DNA
A;Residues: 1-314 <EHR>
A;Cross-references: UNIPROT:P14490; UNIPARC:UPI00001336A6; EMBL:X07371; NID:G46635; PIDN
A;Note: the authors translated the codon GAT for residue 64 as Lys, GCA for residue 174
C;Genetics:
A;Gene: repM
A;Genome: plasmid
C;Superfamily: repC protein

Query Match 65.3%; Score 32; DB 2; Length 314;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
:|||||
Db 243 ESLKEQAMVY 252

RESULT 46
A30480
repJ protein - Staphylococcus aureus plasmid pC223
N;Alternate names: replication initiation protein
C;Species: Staphylococcus aureus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: A30480
R;Projan, S.J.; Novick, R.
Plasmid 19, 203-221, 1988
A;Title: Comparative analysis of five related staphylococcal plasmids.
A;Reference number: J70372; MUID:89161003; PMID:2852816
A;Accession: A30480
A;Molecule type: DNA
A;Residues: 1-314 <PRO>
A;Cross-references: UNIPROT:Q52239; UNIPARC:UPI00001783DB
A;Note: the authors translated the codon tta for residue 123 as Ile
C;Genetics:
A;Gene: repJ
A;Genome: plasmid
C;Superfamily: repC protein
F;238-282/Domain: DNA binding #status predicted <DNB>

Query Match 65.3%; Score 32; DB 2; Length 314;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
:|||||
Db 243 ESLKEQAMVY 252

RESULT 47
T49224
SRG1-like protein - Arabidopsis thaliana
N;Alternate names: protein F27H5.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49224
R;Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25018
A;Accession: T49224
A;Status: preliminary
```

A:Molecule type: DNA
A:Residues: 1-316 <RIE>
A:Cross-references: UNIPROT:Q9LV36; UNIPARC:UPI00000A82E0; EMBL:AL163852; GSPDB:GN00061;
A:Experimental source: cultivar Columbia; BAC clone F27H5
C:Genetics:
A:Gene: ATSP:F27H5.80
A:Map position: 3
A:Introns: 84/3; 126/2
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 65.3%; Score 32; DB 2; Length 316;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 123 EPVREKVGY 132

RESULT 48
T04309
acidic ribosomal protein P0 - rice
C:Species: Oryza sativa (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04309
R:Hinara, Y.; Umeda, M.; Hara, C.; Toriyama, K.; Uchimiya, H.
Plant Physiol. 105, 753-754, 1994
A:Title: Nucleotide sequence of a rice acidic ribosomal phosphoprotein P0 cDNA.
A:Reference number: Z15273; MUID:94345013; PMID:8066137
A:Accession: T04309
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-319 <HH>
A:Cross-references: UNIPROT:P41095; UNIPARC:UPI000013422D; EMBL:D21130; NID:9415316; PID
A:Experimental source: anther
C:Genetics:
A:Gene: YK704
C:Superfamily: rat acidic ribosomal protein P0

Query Match 65.3%; Score 32; DB 2; Length 319;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 99 KEVREEVAKY 108

RESULT 49
T03944
acidic ribosomal protein P0 - maize
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T03944
R:Cordts, S.; Loerz, H.; Dresselhaus, T.
submitted to the EMBL Data Library, September 1996
A:Description: A transcript encoding the large subunit acidic ribosomal protein P0 is s
A:Reference number: Z15151
A:Accession: T03944
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-319 <COR>
A:Cross-references: UNIPROT:O24573; UNIPARC:UPI000016E078; EMBL:Y07959; PIDN:CAA69256.1
A:Experimental source: note inbred line A188
C:Genetics:
A:Note: RP-P0
C:Superfamily: rat acidic ribosomal protein P0
C:Keywords: protein biosynthesis; ribosome; RNA binding

Query Match 65.3%; Score 32; DB 2; Length 319;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 99 KEVREEVAKY 108

RESULT 50
T22701
hypothetical protein F55B12.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22701
R:Sims, M.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19602
A:Accession: T22701
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-350 <WIL>
A:Cross-references: UNIPROT:Q93797; UNIPARC:UPI000007B666; EMBL:Z79757; PIDN:CAB02127.1;
A:Experimental source: clone F55B12
C:Genetics:
A:Gene: CESP:F55B12.6
A:Map position: 5
A:Introns: 64/3; 94/1; 180/2; 227/3; 271/3; 331/1

Query Match 65.3%; Score 32; DB 2; Length 350;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 177 ELIRESVMKH 186

Search completed: May 12, 2006, 10:52:04
Job time : 25.2368 Secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-38
Perfect score: 49
Sequence: 1 ESVREQVMKY 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	83.7	458	Q4S6F3_TETNG	Q4S6f3 tetraodon n
2	40	81.6	1164	Q92JN3_HELPJ	Q92jn3 helicobacte
3	39	79.6	1064	Q4SFR7_TETNG	Q4Sfr7 tetraodon n
4	38	77.6	291	ERA_MYCPN	P75210 mycoplasma
5	38	77.6	592	Q52A99_MAGGR	Q52a99 magnaporthe
6	38	77.6	593	Q5CP31_CRYHO	Q5cp31 cryptospori
7	38	77.6	593	Q5CYP8_CRYPV	Q5cyp8 cryptospori
8	38	77.6	2448	Q6JIZ0_MOUSE	Q6jiz0 mus musculu
9	37	75.5	394	Q9LJZ1_ARATH	Q9ljz1 arabidopsis
10	37	75.5	429	IF5_MOUSE	P59325 mus musculu
11	37	75.5	429	IF5_RAT	Q07205 rattus norv
12	37	75.5	430	Q52TE0_CHICK	Q52te0 gallus gall
13	37	75.5	431	IF5_HUMAN	P55010 homo sapien
14	37	75.5	431	Q53XB3_HUMAN	Q53xb3 homo sapien
15	37	75.5	431	Q5R4L0_PONPY	Q5r4l0 pongo pygma
16	37	75.5	892	Q7QP19_GIALA	Q7qp19 giardia lam
17	37	75.5	1132	Q815C3_PLAF7	Q815c3 plasmodium
18	36	73.5	150	RS13_YEAST	P05756 saccharomyc
19	36	73.5	151	Q6FT65_CANGA	Q6ft65 candida gla
20	36	73.5	160	YBIA_ECOLI	P30176 escherichia
21	36	73.5	160	Q83LJ8_SHIFL	Q83lj8 shigella fl
22	36	73.5	187	Q50H29_BOVIN	Q50h29 bos taurus
23	36	73.5	193	Q4P3L0_USTWA	Q4p3l0 ustilago ma
24	36	73.5	242	Q62DC6_BURMA	Q62dc6 burkholderi
25	36	73.5	245	Q63QH4_BURPS	Q63qh4 burkholderi
26	36	73.5	418	Q6A7W0_PROAC	Q6a7w0 propionibac
27	36	73.5	499	Q9FH67_ARATH	Q9fh67 arabidopsis
28	36	73.5	523	Q9LIG9_ARATH	Q9lig9 arabidopsis
29	36	73.5	671	Q529X0_MAGGR	Q529x0 magnaporthe
30	36	73.5	829	Q51V60_MAGGR	Q51v60 magnaporthe
31	36	73.5	970	Q5CWC1_CRYPV	Q5cwc1 cryptospori

32	36	73.5	998	Q59FY4_HUMAN	Q59fy4 homo sapien
33	36	73.5	1182	Q4SCU4_TETNG	Q4scu4 tetraodon n
34	36	73.5	1233	Q5F363_CHICK	Q5f363 gallus gall
35	36	73.5	1234	JARD2_MOUSE	Q62312 mus musculu
36	36	73.5	1499	Q86L28_SCHMA	Q86l28 schistosoma
37	36	73.5	1847	Q81DM0_PLAF7	Q81dm0 plasmodium
38	36	73.5	2324	COAC_CHICK	P11029 gallus gall
39	36	73.5	2345	COAL_RAT	P11497 rattus norv
40	36	73.5	2345	Q5SWU9_MOUSE	Q5swu9 mus musculu
41	36	73.5	2345	Q6JIZ1_MOUSE	Q6jiz1 mus musculu
42	36	73.5	2346	COAL_HUMAN	Q91ts3 bos taurus
43	36	73.5	2346	COAL_SHEEP	Q13085 homo sapien
44	36	73.5	2346	COAL_HUMAN	Q28559 ovis aries
45	36	73.5	2346	Q7Z561_HUMAN	Q7z561 homo sapien
46	36	73.5	2346	Q6XDA8_HUMAN	Q6xda8 homo sapien
47	35	71.4	74	Q9BS18_HUMAN	Q9bs18 pongo pygma
48	35	71.4	74	Q5RBV4_PONPY	Q5rbv4 pongo pygma
49	35	71.4	74	Q8R034_MOUSE	Q8r034 mus musculu
50	35	71.4	74	Q4SKC6_TETNG	Q4skc6 tetraodon n
51	35	71.4	82	Q5XZ64_BORGA	Q5xz64 borrelia ga
52	35	71.4	127	Q87J10_BORBU	Q87j10 borrelia bu
53	35	71.4	140	Q4RCQ3_TETNG	Q4rcq3 tetraodon n
54	35	71.4	193	Q6MT64_MYCMS	Q6mt64 mycoplasma
55	35	71.4	193	Q5ZKG7_CHICK	Q5zkg7 gallus gall
56	35	71.4	197	Q8MVN6_9ASCI	Q8mvn6 boltenia vi
57	35	71.4	215	Q9HKP2_THEAC	Q9hkp2 thermoplasma
58	35	71.4	217	Q97B50_THEVO	Q97b50 thermoplasma
59	35	71.4	248	Q5ZM54_CHICK	Q5zm54 gallus gall
60	35	71.4	424	Q9R2Y2_BORBU	Q9r2y2 borrelia bu
61	35	71.4	424	Q9S0C6_BORBU	Q9s0c6 borrelia bu
62	35	71.4	424	Q9S0I6_BORBU	Q9s0i6 borrelia bu
63	35	71.4	424	Q9S049_BORBU	Q9s049 borrelia bu
64	35	71.4	429	Q9L9P1_BORBU	Q9l9p1 borrelia bu
65	35	71.4	429	Q7ZWA9_BRARE	Q7zwa9 brachydanio
66	35	71.4	431	Q9S000_BORBU	Q9s000 borrelia bu
67	35	71.4	431	Q9S0F6_BORBU	Q9s0f6 borrelia bu
68	35	71.4	436	Q5XYW7_BORGA	Q5xyw7 borrelia ga
69	35	71.4	444	Q9S093_BORBU	Q9s093 borrelia bu
70	35	71.4	446	Q5ZMN6_CHICK	Q5zmn6 gallus gall
71	35	71.4	486	Q9WZ51_THEMEA	Q9wz51 thermotoga
72	35	71.4	526	Q9LIG8_ARATH	Q9lig8 arabidopsis
73	35	71.4	596	Q76DQ7_9PHYC	Q76dq7 chlorella v
74	35	71.4	653	Q5FRB3_CANGA	Q5frb3 candida gla
75	35	71.4	698	Q51X41_MAGGR	Q51x41 magnaporthe
76	35	71.4	753	Q4N7W0_THEPA	Q4n7w0 theileria p
77	35	71.4	810	REB1_YEAST	P21538 saccharomyc
78	35	71.4	869	Q6OSF7_CAEBR	Q6osf7 caenorhabdi
79	35	71.4	1003	Q8IAC0_HALRO	Q8iac0 halocynthia
80	35	71.4	1271	P94460_BACSU	P94460 bacillus su
81	35	71.4	1272	Q799M0_BACSU	Q799m0 bacillus su
82	35	71.4	1279	Q31827_BACSU	Q31827 bacillus su
83	35	71.4	1317	Q54X69_DICDI	Q54x69 dictyosteli
84	35	71.4	2456	Q70151_RAT	Q70151 rattus norv
85	34	69.4	99	Q6LNQ3_PROPR	Q6lnq3 photobacter
86	34	69.4	132	Q4XH90_PLACH	Q4xh90 plasmodium
87	34	69.4	159	Q6YD94_9FABA	Q6y940 arachis ste
88	34	69.4	159	Q6YDF0_9FABA	Q6ydf0 arachis car
89	34	69.4	175	Q81EH3_BACCR	Q81eh3 bacillus ce
90	34	69.4	175	Q63CF3_BACCR	Q63cf3 bacillus ce
91	34	69.4	204	Q8SSK8_ENCCU	Q8ssk8 encephalito
92	34	69.4	212	Q9KRL5_VIBCH	Q9krl5 vibrio chol
93	34	69.4	219	Q8NLY1_CORGL	Q8nly1 corynebacte
94	34	69.4	219	BUN1_DROME	Q24522 drosophila
95	34	69.4	272	Q98QV3_MYCPU	Q98qv3 mycoplasma
96	34	69.4	287	Q48615_9LACT	Q48615 lactococcus
97	34	69.4	357	Q8Z110_SALTI	Q8z110 salmonella
98	34	69.4	383	Q59RJ2_CANAL	Q59rj2 candida alb
99	34	69.4	407	Q5R5T5_PONPY	Q5rst5 pongo pygma
100	34	69.4	412	Q8QGW4_BRARE	Q8qgw4 brachydanio
101	34	69.4	416	Q6MM69_BDEBA	Q6mm69 bdellovibri
102	34	69.4	450	Q4R7L2_MACFA	Q4r7l2 macaca fasc
103	34	69.4	455	Q98RI8_MYCPU	Q98ri8 mycoplasma
104	34	69.4	481	Q584W0_9TRYP	Q584w0 trypanosoma

105	34	69.4	504	2	Q896Y2	Clostridium	178	33	67.3	461	2	Q5A441	CANAL	Q5A441	candida alb
106	34	69.4	513	2	Q8LEJ8	arabidopsis	179	33	67.3	469	2	Q4ZXX9	PSEY	Q4ZXX9	pseudomonas
107	34	69.4	513	2	Q8LJY8	arabidopsis	180	33	67.3	490	2	Q9CEV0	LACIA	Q9CEV0	lactococcus
108	34	69.4	516	2	Q8BU60	debaromyces	181	33	67.3	497	2	Q7RLR0	PLAYO	Q7RLR0	plasmidium
109	34	69.4	519	2	Q8A8M6	protonibac	182	33	67.3	505	2	Q89IM1	CLOTE	Q89IM1	clostridium
110	34	69.4	519	2	Q8A8M6	pyrococcus	183	33	67.3	512	2	Q9XV10	CAEEL	Q9XV10	caenorhabdi
111	34	69.4	523	1	Q8A8M6	pyrococcus	184	33	67.3	514	2	Q9ZVD7	ARATH	Q9ZVD7	arabidopsis
112	34	69.4	525	2	Q84421	parametium	185	33	67.3	542	2	Q6WGM5	PARA	Q6WGM5	mossman vir
113	34	69.4	752	2	Q7Q3A1	anopheles g	186	33	67.3	514	2	Q55VJ1	CRYNE	Q55VJ1	cryptococcus
114	34	69.4	811	1	PLSB	haemophilus	187	33	67.3	561	2	Q5KPB8	CRYNE	Q5KPB8	cryptococcus
115	34	69.4	825	1	RCAL	saccharomyc	188	33	67.3	561	2	Q6MQB2	DEBEA	Q6MQB2	bdellovibri
116	34	69.4	825	1	Q6B111	saccharomyc	189	33	67.3	567	2	Q5ACH4	CANAL	Q5ACH4	candida alb
117	34	69.4	834	2	Q8SXRO	drosophila	190	33	67.3	569	1	MALT	CANAL	Q8Sach4	candida alb
118	34	69.4	858	2	Q4G170	human	191	33	67.3	570	2	MUS81	SCHPO	P87231	schizosacch
119	34	69.4	861	2	Q4HB46	deinococcus	192	33	67.3	572	1	REB1	KLULA	Q85950	kluyveromyc
120	34	69.4	906	2	Q25646	plasmidium	193	33	67.3	595	1	Q8JRU0	9BACU	Q8JRU0	phthorimaea
121	34	69.4	916	1	CADH4	human	194	33	67.3	609	2	Q5WXV3	LEGPL	Q5WXV3	legionella
122	34	69.4	924	2	Q9NA99	caenorhabdi	195	33	67.3	673	2	Q5X6B6	LEGPA	Q5X6B6	legionella
123	34	69.4	940	1	UVRA	VIBVU	196	33	67.3	683	2	Q5ZWX6	LEGPH	Q5ZWX6	legionella
124	34	69.4	940	1	UVRA	VIBVU	197	33	67.3	751	2	Q9FMN1	ARATH	Q9FMN1	arabidopsis
125	34	69.4	949	2	Q8EBZ1	shevanella	198	33	67.3	776	2	Q94LH1	ORYSA	Q94LH1	oryza sativ
126	34	69.4	960	2	Q8JRM5	cherry rasp	199	33	67.3	784	2	Q5GAP5	MAIZE	Q5GAP5	zea mays (m
127	34	69.4	960	2	Q8EWG8	cherry rasp	200	33	67.3	798	2	Q75WV8	BRABE	Q75WV8	branchiosto
128	34	69.4	968	2	Q5P418	azocarcus sp	201	33	67.3	817	2	Q8LHO6	ORYSA	Q8LHO6	oryza sativ
129	34	69.4	984	2	Q4Z4G1	plasmidium	202	33	67.3	836	2	Q7YVH4	CRYPV	Q7YVH4	cryptospori
130	34	69.4	986	2	Q4XZ91	plasmidium	203	33	67.3	862	1	RABE1	HUMAN	Q5C276	homo sapien
131	34	69.4	1048	2	Q7RS02	plasmidium	204	33	67.3	896	2	Q5CX78	CRYPV	Q5CX78	cryptospori
132	34	69.4	1070	2	Q8IB03	plasmidium	205	33	67.3	921	2	Q7NBT1	MYCGA	Q7NBT1	mycoplasma
133	34	69.4	1689	2	Q59GJ9	human	206	33	67.3	921	2	Q8ZJY3	SALTY	Q8ZJY3	salmonella
134	34	69.4	1838	2	Q8ID26	plasmidium	207	33	67.3	957	2	Q6CL70	KLULA	Q6CL70	kluyveromyc
135	34	69.4	2458	2	Q6KE87	human	208	33	67.3	1012	2	Q580R3	9TRYP	Q580R3	trypanosoma
136	34	69.4	2458	2	Q6TY48	human	209	33	67.3	1030	2	Q6FLN2	CANGA	Q6FLN2	candida gla
137	33	67.3	2483	1	COA2	human	210	33	67.3	1153	2	Q97GV2	CLOAB	Q97GV2	clostridium
138	33	67.3	72	2	Q5QF4	planktothri	211	33	67.3	1202	2	Q21303	CAEEL	Q21303	caenorhabdi
139	33	67.3	72	2	Q5QQQ6	planktothri	212	33	67.3	1207	2	Q60TI2	CAEEL	Q60TI2	caenorhabdi
140	33	67.3	72	2	Q5QQQ8	planktothri	213	33	67.3	1287	2	Q7PDW0	PLAYO	Q7PDW0	plasmidium
141	33	67.3	72	2	Q5QQR2	oscillatori	214	33	67.3	1451	2	Q5KHJ6	CRYNE	Q5KHJ6	cryptococcus
142	33	67.3	72	2	Q5QQR6	oscillatori	215	33	67.3	1462	2	Q55UU2	CRYNE	Q55UU2	cryptococcus
143	33	67.3	72	2	Q5QQR0	oscillatori	216	33	67.3	1873	1	CAC15	RABIT	P07293	oryctolagus
144	33	67.3	72	2	Q5QQR2	oscillatori	217	33	67.3	2365	2	Q4RSU6	TETNG	Q4RSU6	tetradon n
145	33	67.3	111	2	Q9WMG1	topogratov	218	33	67.3	3488	2	Q8G986	OSCAG	Q8G986	oscillatori
146	33	67.3	133	2	Q4RJ93	tetraodon n	219	32	65.3	69	2	Q72UT4	LEPIC	Q72UT4	leptospira
147	33	67.3	136	2	Q87H2	neurospora	220	32	65.3	103	2	Q6LNF8	PHOPR	Q6LNF8	photobacter
148	33	67.3	159	2	Q5JE92	pyrococcus	221	32	65.3	134	1	CALBP	ENTHI	P38505	entamoeba h
149	33	67.3	163	2	Q704P4	bos taurus	222	32	65.3	134	2	Q517B3	ENTHI	Q517B3	entamoeba h
150	33	67.3	179	2	Q7W67	sulfolobus	223	32	65.3	138	2	Q23434	CAEEL	Q23434	caenorhabdi
151	33	67.3	219	2	Q4H9X6	deinococcus	224	32	65.3	165	2	Q7N9G4	PHOLL	Q7N9G4	photorhabdi
152	33	67.3	228	2	Q5SGC2	bacillus li	225	32	65.3	170	2	Q9P4B4	LETUV	Q9P4B4	letharia vu
153	33	67.3	247	2	Q7RYB7	neurospora	226	32	65.3	184	2	Q8XL92	CLOPE	Q8XL92	clostridium
154	33	67.3	270	2	Q8A0A3	bacteroides	227	32	65.3	184	2	Q848T4	ENTFC	Q848T4	enterococcus
155	33	67.3	273	2	Q4Z306	plasmidium	228	32	65.3	184	2	Q5M8K8	XENTR	Q5M8K8	xenopus tro
156	33	67.3	278	2	Q41851	zea mays (m	229	32	65.3	193	2	Q5C190	SCHJA	Q5C190	schistosoma
157	33	67.3	280	2	Q5LUG3	bacteroides	230	32	65.3	197	2	Q678E9	9VIRU	Q678E9	lymphocyati
158	33	67.3	336	2	Q64XP6	bacteroides	231	32	65.3	198	2	Q6Q912	9GAMM	Q6Q912	uncultured
159	33	67.3	354	2	Q7T101	xenopus lae	232	32	65.3	200	2	Q6P2S7	HUMAN	Q6P2S7	homo sapien
160	33	67.3	354	2	Q4QR41	xenopus lae	233	32	65.3	203	2	Q8LNX8	ZINEL	Q8LNX8	zinnia eleg
161	33	67.3	355	2	Q28382	archaeoglob	234	32	65.3	205	2	Q5T829	HUMAN	Q5T829	homo sapien
162	33	67.3	358	1	ALF	SCHPO	235	32	65.3	214	2	Q5E243	VIBF1	Q5E243	vibrio fisc
163	33	67.3	366	2	Q50PH3	entamoeba h	236	32	65.3	217	2	Q97DY0	CLOAB	Q97DY0	clostridium
164	33	67.3	377	2	Q50V88	entamoeba h	237	32	65.3	222	1	Y175	METJA	Y175	metanococc
165	33	67.3	384	2	Q6DG41	brachydanio	238	32	65.3	224	2	Q993S6	9VIRU	Q993S6	banana mild
166	33	67.3	384	2	Q54138	dictyosteli	239	32	65.3	227	2	Q516A0	ENTHI	Q516A0	entamoeba h
167	33	67.3	392	2	Q8YUP3	anabaena sp	240	32	65.3	227	2	Q8DKV2	SYNEL	Q8DKV2	synecchococ
168	33	67.3	398	2	Q5QVZ0	idiomarina	241	32	65.3	236	2	Q600A4	MYCHY	Q600A4	mycoplasma
169	33	67.3	400	1	ACKA	PHOLL	242	32	65.3	240	2	Q5YU16	NOCFA	Q5YU16	nocardia fa
170	33	67.3	377	2	Q6DG41	brachydanio	243	32	65.3	247	2	Q86D97	SCHMA	Q86D97	schistosoma
171	33	67.3	384	2	Q54138	dictyosteli	244	32	65.3	251	2	Q84VC9	ORYSA	Q84VC9	oryza sativ
172	33	67.3	392	2	Q8YUP3	anabaena sp	245	32	65.3	252	2	Q4X1T0	ASPTU	Q4X1T0	aspergillus
173	33	67.3	398	2	Q5QVZ0	idiomarina	246	32	65.3	260	2	Q7ZY55	XENLA	Q7ZY55	xenopus lae
174	33	67.3	400	1	ACKA	PHOLL	247	32	65.3	260	2	Q7SZAI	XENLA	Q7SZAI	xenopus lae
175	33	67.3	413	2	Q8TMS1	metac	248	32	65.3	261	2	Q6DDCS	XENTR	Q6DDCS	xenopus tro
176	33	67.3	431	2	Q6IBU0	homo sapien	249	32	65.3	264	2	Q8AW09	BRARE	Q8AW09	brachydanio
177	33	67.3	444	2	Q8A735	bacteroides	250	32	65.3	266	2	Q51SN5	MACFA	Q51SN5	macaca faec

251	32	65.3	275	2	Q4L7A8_STAHLJ	Q4L7A8 staphylococ	324	721	2	Q9PA43_XYLFA	Q9pa43 xyella fas
252	32	65.3	286	2	Q7ZZ43_BRARE	P7zz43 brachydanio	325	732	2	Q5LAR7_BACFN	Q5lar7 bacteroides
253	32	65.3	301	1	SC02_YEAST	P38072 saccharomyc	326	732	2	Q64R64_BACFR	Q64r64 bacteroides
254	32	65.3	303	2	Q74I79_LACJO	Q74i79 lactobacill	327	754	2	Q8BRK4_MOUSE	Q8brk4 mus musculu
255	32	65.3	304	2	Q4HN33_CAMLA	Q4hn33 campylobact	328	759	2	Q9JZ50_NEIMB	Q9jz50 neisseria m
256	32	65.3	314	1	REPM_STAUA	P14490 staphylococ	329	782	2	Q74QJ8_YERPE	Q74qj8 yersinia pe
257	32	65.3	314	2	Q52239_plasmid pc2	Q52239 plasmid pc2	330	782	2	Q8CKW7_YERPE	Q8ckw7 yersinia pe
258	32	65.3	314	2	Q6V0V3_STAUA	Q6v0v3 staphylococ	331	782	2	Q8ZHC8_YERPE	Q8zhc8 yersinia pe
259	32	65.3	316	2	Q9LX36_ARATH	Q9lx36 arabidopsis	332	782	2	Q8ZG45_YERPE	Q8zg45 yersinia ps
260	32	65.3	318	1	RLA0_MAZE	Q24573 zea mays (m	333	782	2	Q666K9_YERPS	Q666k9 yersinia ps
261	32	65.3	319	1	RLA0_ORYSA	P41095 oryza sativ	334	799	1	CADH8_HUMAN	P52286 homo sapien
262	32	65.3	319	2	Q7EZR2_ORYSA	Q7ezr2 oryza sativ	335	799	1	CADH8_MOUSE	P97291 mus musculu
263	32	65.3	330	2	Q5KVB0_GEOKA	Q5kvb0 geobacillus	336	799	1	CADH8_RAT	O54800 rattus norv
264	32	65.3	337	2	Q54EC9_DICDI	Q54ec9 dictyosteli	337	824	2	Q8FED2_ECOL6	Q8fed2 escherichia
265	32	65.3	339	2	Q5CH22_YARLW	Q5ch22 yarrowia li	338	832	2	Q5VT46_HUMAN	Q5vt46 homo sapien
266	32	65.3	349	2	Q5GHP0_LACRH	Q5ghf0 lactobacill	339	832	2	Q66CB6_YERPS	Q66cb6 yersinia ps
267	32	65.3	350	2	Q93797_CABEL	Q93797 caenorhabdi	340	841	2	Q7N257_PHOLL	Q7n257 photorhabdu
268	32	65.3	350	2	Q8PER8_XANAC	Q8per8 xanthomonas	341	842	2	Q7N245_PHOLL	Q7n245 photorhabdu
269	32	65.3	354	2	Q9QZA9_MOUSE	Q9qza9 mus musculu	342	854	1	DISC1_HUMAN	Q9nris homo sapien
270	32	65.3	359	2	Q5I0L6_RAT	Q5i0l6 rattus norv	343	854	1	Q5KX82_GEOKA	Q5kx82 geobacillus
271	32	65.3	359	2	Q9LVY6_MOUSE	Q9lvv6 mus musculu	344	857	2	Q5QWH4_IDILO	Q5qwh4 idiomarina
272	32	65.3	375	2	Q8F3S2_LEPIN	Q8f3s2 leptospira	345	872	2	Q74I88_LACJO	Q74i88 lactobacill
273	32	65.3	375	2	Q72RX4_LEPIC	Q72rx4 leptospira	346	880	2	Q66CB3_YERPS	Q66cb3 yersinia ps
274	32	65.3	392	1	OGG1_YEAST	P53397 saccharomyc	347	900	2	Q6B390_9PROT	Q6b390 uncultured
275	32	65.3	392	2	Q6GN39_XENLA	Q6gn39 xenopus lae	348	910	2	Q8B390_9PROT	Q8b390 anabaena sp
276	32	65.3	394	2	Q64CZ1_9ARCH	Q64cz1 uncultured	349	930	2	Q4YMS8_ANASP	Q4yus2 theileria a
277	32	65.3	403	1	METK_XANAC	Q8pp75 xanthomonas	350	937	2	Q4YU20_THEAN	Q8kuw5 vibrio chol
278	32	65.3	403	1	METK_XANCP	Q8pch3 xanthomonas	351	940	1	UVRA_VIBCH	Q87la0 vibrio para
279	32	65.3	403	1	METK_XYLFA	Q8p9b0 xyella fas	352	940	1	UVRA_VIBPA	Q87la0 vibrio para
280	32	65.3	403	1	METK_XYLFT	Q8p9b0 xyella fas	353	952	2	Q8ZVJ8_PYRAE	Q8zvj8 pyrobaculum
281	32	65.3	403	2	Q5GW76_XANCP	Q5gw76 xanthomonas	354	969	2	Q5CJ01_CRYHO	Q5cjo1 cryptospori
282	32	65.3	403	2	Q4UR08_XANCP	Q4ur08 xanthomonas	355	986	2	Q900J9_9RETR	Q900j9 ovine lenti
283	32	65.3	403	2	Q88GQ1_PSEPK	Q88gq1 pseudomonas	356	1045	2	P90884_CAEEL	P90884 caenorhabdi
284	32	65.3	403	2	Q97LQ3_CLOAB	Q97lq3 clostridium	357	1050	2	Q7Q1Y1_ANOGA	Q7q1y1 anopheles g
285	32	65.3	410	2	Q9YCU8_AERPE	Q9ycul aeropyrum p	358	1057	2	Q61D01_CAEBR	Q61dq1 caenorhabdi
286	32	65.3	410	2	P96482_STRPN	P96482 streptococc	359	1068	2	Q6FNM9_CANGA	Q6fnm9 candida gla
287	32	65.3	410	2	Q4K386_STRPN	Q4k386 streptococc	360	1136	2	Q5AT35_EMENI	Q5at35 aspergillus
288	32	65.3	412	2	Q9X9A9_STRPN	Q9x9a9 streptococc	361	1167	2	Q7FPM0_ANOGA	Q7fpm0 anopheles g
289	32	65.3	412	2	Q9AH99_STRPN	Q9ah99 streptococc	362	1246	1	Q7RWR6_NEUCR	Q7rwr6 neurospora
290	32	65.3	428	2	Q4JAY2_SULAC	Q4jay2 sulfolobus	363	1267	2	JARD2_HUMAN	Q92833 homo sapien
291	32	65.3	433	2	Q4NCL8_9MICC	Q4nc18 arthrobacte	364	1571	2	Q4N6Y1_THEPA	Q4n6y1 theileria p
292	32	65.3	444	1	FIBG_BOVIN	P12799 bos taurus	365	1620	1	Q59ZR2_MOUSE	Q59zr2 mus musculu
293	32	65.3	445	2	Q97J59_CLOAB	Q97j59 clostridium	366	1620	1	HECD1_HUMAN	Q9ult8 homo sapien
294	32	65.3	445	2	Q9KDG8_BACHD	Q9kdg8 bacillus ha	367	1899	2	Q4QEU6_LEIMA	Q4qeu6 leishmania
295	32	65.3	451	2	Q83DH6_COXBU	Q83dh6 coxiella bu	368	2148	2	Q6QTA9_9VIRU	Q6qta9 pumala vir
296	32	65.3	460	2	Q6N8G4_RHOPA	Q6n8g4 rhodospheudo	369	2156	1	RRPL_PUMH	P27176 pumala vir
297	32	65.3	468	2	Q9LCC6_9BACI	Q9lcc6 bacillus sp	370	2271	2	Q4I878_GIBZE	Q4i878 gibberella
298	32	65.3	476	2	Q4H579_DEBIO	Q4h579 deinococcus	371	2612	2	Q86VJ1_HUMAN	Q86vj1 homo sapien
299	32	65.3	476	2	Q6BUY1_DEBHA	Q6buy1 debaryomyce	372	2771	2	Q4Y148_PLACH	Q4y148 plasmodium
300	32	65.3	488	2	Q4FWX7_TETNG	Q4fwx7 tetraodon n	373	56	2	Q95J69_PANTR	Q95j69 pan troglod
301	32	65.3	489	1	Y052_METJA	Q57557 methanococc	374	60	2	Q8F177_ECOL6	Q8f177 escherichia
302	32	65.3	496	2	Q4QCG7_LEIMA	Q4qcg7 leishmania	375	63	2	P72432_STRSQ	P72432 streptococc
303	32	65.3	501	2	Q9WYF3_THEMA	Q9wyf3 thermotoga	376	88	2	Q6YSC4_ORYSA	Q6yssc4 oryza sativ
304	32	65.3	526	2	Q68F97_XENTR	Q68f97 xenopus tro	377	91	1	RL34_PYRAB	Q9uzj7 pyrococcus
305	32	65.3	530	1	WC2_NEUCR	P78714 neurospora	378	91	1	RL34_PYRAB	Q74006 pyrococcus
306	32	65.3	530	2	Q7M8P0_WOLSU	Q7m8p0 wolinnella s	379	95	2	Q4FLQ6_9RICK	Q4flq6 candidatus
307	32	65.3	531	2	Q67J74_SYMTH	Q67j74 symbiobacte	380	96	1	Y3713_CLOAB	P34159 clostridium
308	32	65.3	538	2	Q7RVJ5_NEUCR	Q7rvj5 neurospora	381	96	1	Q56BT3_9CAUD	Q56bt3 enterobacte
309	32	65.3	548	2	Q7PYS5_ANOGA	Q7pys5 anopheles g	382	102	2	Q6YSC3_ORYSA	Q6ysc3 oryza sativ
310	32	65.3	550	2	Q60WT6_CAEBR	Q60wt6 caenorhabdi	383	110	2	Q6MBX6_PAKUM	Q6mbx6 parachlamyd
311	32	65.3	550	2	O22880_ARATH	O22880 arabidopsis	384	113	2	Q6HNE2_BACHK	Q6hne2 bacillus th
312	32	65.3	570	2	Q12457_YEAST	Q12457 saccharomyc	385	113	2	Q81V31_BACAN	Q81v31 bacillus an
313	32	65.3	573	2	Q4SUC2_TETNG	Q4suc2 tetraodon n	386	116	2	Q63FX4_BACCC	Q63fx4 bacillus ce
314	32	65.3	578	2	Q60J20_CAEBR	Q60j20 caenorhabdi	387	117	2	Q8DSI8_STRMU	Q8dsi8 streptococc
315	32	65.3	589	2	Q501D5_ARATH	Q501d5 arabidopsis	388	118	1	RS13_SHEON	Q8ek48 shewanella
316	32	65.3	595	2	Q53MB3_ORYSA	Q53mb3 oryza sativ	389	120	1	RS13_NEIMA	P66385 neisseria m
317	32	65.3	658	1	Q9AT32_DAUCA	Q9at32 daucus caro	390	120	2	Q5F5U9_NEIG1	P66386 neisseria g
318	32	65.3	678	1	VPS1_SCHPO	Q9urz5 schizosacch	391	120	2	Q7PAI3_RICSI	Q7pai3 rickettsia
319	32	65.3	685	2	Q5LUQ7_MAGOR	Q5luq7 magnaporthe	392	127	2	Q9ZE94_RICPR	Q9ze94 rickettsia
320	32	65.3	711	2	Q8I411_CABEL	Q8i411 caenorhabdi	393	127	2	HVA22_HORVU	Q07764 hordeum vul
321	32	65.3	714	2	Q7SX85_BRARE	Q7sx85 brachydanio	394	130	1	Q7XAP6_WHEATY	Q7xap6 triticum ae
322	32	65.3	716	2	Q8C449_MOUSE	Q8c449 mus musculu	395	131	2	Q68XS6_RICTY	Q68xs6 rickettsia
323	32	65.3	721	2	Q87A00_XYLFT	Q87a00 xyella fas	396	131	2		

397	31	63.3	131	2	Q90276	boulengerel	470	31	63.3	274	2	Q4MLJ8	BACCE	Q4mlj8	bacillus ce
398	31	63.3	133	2	Q774U3	cotton leaf	471	31	63.3	274	2	Q6HH25	BACHK	Q6hh25	bacillus th
399	31	63.3	133	2	Q774U7	cotton leaf	472	31	63.3	274	2	Q735Q4	BACCL	Q735q4	bacillus ce
400	31	63.3	133	2	Q77M72	cotton leaf	473	31	63.3	274	2	Q813F6	BACCR	Q813f6	bacillus ce
401	31	63.3	133	2	Q8UYG4	cotton leaf	474	31	63.3	274	2	Q81NV3	BACAN	Q81nv3	bacillus an
402	31	63.3	133	2	Q8V0H7	cotton leaf	475	31	63.3	274	2	Q639P3	BACCCZ	Q639p3	bacillus ce
403	31	63.3	133	2	Q91N50	cotton leaf	476	31	63.3	275	2	Q9MSA7	MAIZE	Q9msa7	zea mays (m
404	31	63.3	134	2	Q5L4H1	east africa	477	31	63.3	282	2	Q6A0A8	DESPS	Q6a0a8	desulfotale
405	31	63.3	136	2	Q660X8	borgia	478	31	63.3	283	2	Q9LQ05	ARATH	Q9lq05	arabidopsis
406	31	63.3	139	2	Q92J30	rickettsia	479	31	63.3	283	2	Q8LG94	ARATH	Q8lg94	arabidopsis
407	31	63.3	141	2	Q592R8	lymaea sta	480	31	63.3	284	1	Y1834	THETN	Q8r828	thermoanaer
408	31	63.3	149	2	Q4S576	TETNG	481	31	63.3	284	2	Q8GXB5	ARATH	Q8gxb5	arabidopsis
409	31	63.3	150	2	Q8RFL1	FUSN	482	31	63.3	291	2	Q9UZ17	PYRAB	Q9uz17	pyrococcus
410	31	63.3	151	2	Q75011	ASHGO	483	31	63.3	291	2	Q9FPC8	ARATH	Q9fpc8	arabidopsis
411	31	63.3	153	2	Q8ERJ5	OCEI	484	31	63.3	292	2	Q97V56	SULSO	Q97v56	eulfolobus
412	31	63.3	158	1	GRDA2	EUBAC	485	31	63.3	292	2	Q8Y536	LISMO	Q8y536	listeria mo
413	31	63.3	164	2	Q9L770	STREQ	486	31	63.3	292	2	Q71XC5	LISMF	Q71xc5	listeria mo
414	31	63.3	165	2	Q8E286	STRA5	487	31	63.3	293	2	Q8TL92	METAC	Q8tl92	methanosarc
415	31	63.3	165	2	Q8E7P4	STRA3	488	31	63.3	294	1	HD1	BRANA	P46606	brassica na
416	31	63.3	165	2	Q5X4V8	LEGPA	489	31	63.3	295	1	ST1E1	BOVIN	P19217	bos taurus
417	31	63.3	166	2	Q912D7	PSEAE	490	31	63.3	300	2	Q7MB45	VIBVY	Q7me45	vibrio vuln
418	31	63.3	174	2	Q720C0	LISMF	491	31	63.3	301	2	Q4SF76	TETNG	Q4sf76	tetraodon n
419	31	63.3	175	2	Q5ECC7	9PARA	492	31	63.3	311	1	REP	STAAU	P03065	staphylococ
420	31	63.3	175	2	Q5ECD4	9PARA	493	31	63.3	314	1	REP	STAAU	P08115	staphylococ
421	31	63.3	176	2	Q9LAV3	9STRE	494	31	63.3	315	1	ALDR	HUMAN	P15121	homo sapien
422	31	63.3	181	2	Q5ECC8	9PARA	495	31	63.3	316	2	Q6ICP2	HUMAN	Q6icp2	homo sapien
423	31	63.3	181	2	Q5ECD5	9PARA	496	31	63.3	316	2	Q6FGA4	HUMAN	Q6fga4	homo sapien
424	31	63.3	187	2	Q74HV5	LACJO	497	31	63.3	316	2	Q5U031	HUMAN	Q5u031	homo sapien
425	31	63.3	188	2	Q9S171	ARATH	498	31	63.3	316	2	Q5RAB3	PONPY	Q5rab3	pongo pygma
426	31	63.3	189	2	Q5HX51	CAMJR	499	31	63.3	318	2	Q8THQ3	METAC	Q8thq3	methanosarc
427	31	63.3	189	2	Q9PJ12	CAMJE	500	31	63.3	323	2	Q4QJD6	LEIMA	Q4qjd6	leishmania
428	31	63.3	190	2	Q7R7T9	PLAYO	501	31	63.3	326	2	Q5X2H8	LEGPA	Q5x2h8	legionella
429	31	63.3	199	2	Q6F265	MESFL	502	31	63.3	326	2	Q6NRJ5	XENLA	Q6nrj5	xenopus lae
430	31	63.3	200	2	Q8Q7V3	9HIV1	503	31	63.3	330	2	Q6SI33	9BACT	Q6si33	uncultured
431	31	63.3	204	2	Q5ECC9	9PARA	504	31	63.3	330	2	Q8CZA7	STRR6	Q8cza7	streptococc
432	31	63.3	204	2	Q5ECD6	9PARA	505	31	63.3	334	2	Q6BM70	DEBHA	Q6bm70	debaryomyce
433	31	63.3	206	2	Q8CY81	STRER6	506	31	63.3	335	2	Q81MY1	BACAN	Q81my1	bacillus an
434	31	63.3	206	2	Q7N8G6	STRPN	507	31	63.3	341	2	Q9Z746	CHLPN	Q9z746	chlamydia p
435	31	63.3	206	2	Q7ZWN7	XENLA	508	31	63.3	349	2	Q6GJ96	XENLA	Q6gj96	xenopus lae
436	31	63.3	209	2	Q8R772	THETN	509	31	63.3	349	2	Q66JL0	XENTR	Q66jl0	xenopus tro
437	31	63.3	211	2	Q9EUP5	THET8	510	31	63.3	350	2	Q5H6B4	XANOR	Q5h6b4	xanthomonas
438	31	63.3	211	2	Q5SKT0	THET8	511	31	63.3	350	2	Q4UNW3	XANCP	Q4unw3	xanthomonas
439	31	63.3	211	2	Q72L68	THET2	512	31	63.3	350	2	Q8P3B1	XANCP	Q8p3b1	xanthomonas
440	31	63.3	212	2	Q8T924	TROOR	513	31	63.3	352	2	Q9IU81	9HIV1	Q9iu81	human immun
441	31	63.3	212	2	Q73J71	TREDE	514	31	63.3	354	2	Q51SX7	ENTHI	Q51sx7	entamoeba h
442	31	63.3	213	1	KTHY	MYCGA	515	31	63.3	361	2	Q5GFV9	9HIV1	Q5gfv9	human immun
443	31	63.3	214	1	C.SEND6		516	31	63.3	362	2	Q5OMW3	ENTHI	Q5omw3	entamoeba h
444	31	63.3	214	1	C.SEND6		517	31	63.3	363	2	Q9D511	MOUSE	Q9d511	mus musculus
445	31	63.3	214	1	C.SENDP		518	31	63.3	363	2	Q9CWV4	MOUSE	Q9cwv4	mus musculus
446	31	63.3	214	1	C.SENDH		519	31	63.3	366	2	Q61WV9	CAEBR	Q61wv9	caenorhabdi
447	31	63.3	214	1	C.SENDN		520	31	63.3	366	2	Q9F763	BACFR	Q9f763	bacteroides
448	31	63.3	214	1	C.SENDN		521	31	63.3	366	2	Q5LE55	BACFN	Q5le55	bacteroides
449	31	63.3	214	1	C.SENDZ		522	31	63.3	366	2	Q64V97	BACFR	Q64v97	bacteroides
450	31	63.3	215	2	Q5ECD3	9PARA	523	31	63.3	369	2	Q5H2B4	RAT	Q5h2b4	rattus norv
451	31	63.3	215	2	Q5ECD0	9PARA	524	31	63.3	369	2	Q8BXP6	MOUSE	Q8bxf6	mus musculus
452	31	63.3	225	2	Q7YSV7	DROSI	525	31	63.3	369	2	Q9CQT1	MOUSE	Q9cqt1	m mus muscu
453	31	63.3	225	2	Q7YSW6	DROPHILA	526	31	63.3	374	2	Q751Y0	ORYSA	Q751y0	oryza sativ
454	31	63.3	225	2	Q7YSW6	DROSI	527	31	63.3	378	2	Q50LM9	ENTHI	Q50lm9	entamoeba h
455	31	63.3	225	2	Q7YUD0	DROSI	528	31	63.3	378	2	Q8GXU3	ARATH	Q8gxu3	arabidopsis
456	31	63.3	225	2	Q7YUD1	DROSI	529	31	63.3	380	2	Q581Q3	9TRYP	Q581q3	trypanosoma
457	31	63.3	225	2	Q7YUD2	DROSI	530	31	63.3	381	1	METK	AMOPS	Q7wyg5	amoeba prot
458	31	63.3	225	2	Q7YUD3	DROSI	531	31	63.3	384	2	Q84LI9	ORYSA	Q84li9	oryza sativ
459	31	63.3	225	2	Q9BJZ7	DROSE	532	31	63.3	384	2	Q6L8F7	ORYSA	Q6l8f7	oryza sativ
460	31	63.3	225	2	Q9BJZ8	DROMA	533	31	63.3	384	2	Q4MTU8	BACCE	Q4mtu8	bacillus ce
461	31	63.3	231	2	Q5OVX7	ENTHI	534	31	63.3	385	2	Q81DW8	BACCR	Q81dw8	bacillus ce
462	31	63.3	237	2	Q8VY70	ARATH	535	31	63.3	385	2	Q5JLE5	ORYSA	Q5jle5	oryza sativ
463	31	63.3	238	2	Q8T9V9	9DIPT	536	31	63.3	386	2	Q7D6P7	MYCTU	Q7d6p7	mycobacteri
464	31	63.3	239	2	Q8SQ85	ENCOCU	537	31	63.3	386	2	Q33229	MYCTU	Q33229	mycobacteri
465	31	63.3	241	2	Q8T9Z0	DROTK	538	31	63.3	386	2	Q7TY12	MYCBO	Q7ty12	mycobacteri
466	31	63.3	248	2	Q8PTA1	METMA	539	31	63.3	389	2	Q13025	XENLA	Q13025	xenopus lae
467	31	63.3	259	1	IF2C	GALSU	540	31	63.3	391	2	Q8EW51	MYCPE	Q8ew51	mycoplasma
468	31	63.3	260	2	Q59EL5	HUMAN	541	31	63.3	391	2	Q5X3J7	LEGPA	Q5x3j7	legionella
469	31	63.3	268	2	Q9KMK5	VIBCH	542	31	63.3	392	2	Q4ZRX3	PSESY	Q4zrx3	pseudomonas

543	31	63.3	396	1	IMPK_MOUSE	Q7tt16 mus musculus
544	31	63.3	399	1	IMPK_RAT	Q99n14 rattus norv
545	31	63.3	399	2	Q5ZT03_LEGPH	Q5zt03 legionella
546	31	63.3	402	2	Q66J00_XENLA	Q66j00 xenopus lae
547	31	63.3	411	2	Q7PZU6_ANOGA	Q7pzu6 anopheles g
548	31	63.3	413	2	Q5V6Q7_HALMA	Q5v6q7 haloarcula
549	31	63.3	415	2	Q82CK7_YERPE	Q82ck7 yersinia pe
550	31	63.3	415	2	Q668N8_YERPS	Q668n8 yersinia ps
551	31	63.3	416	1	IMPK_HUMAN	Q8nfus homo sapien
552	31	63.3	426	2	Q55F21_DICDI	Q55f21 dictyosteli
553	31	63.3	427	1	TCO2_RAT	Q9r0d6 rattus norv
554	31	63.3	430	1	TCO2_MOUSE	Q89968 mus musculus
555	31	63.3	430	1	YAE9_SCHPO	Q93849 schizosacch
556	31	63.3	430	2	Q5SQ21_MOUSE	Q5sq21 mus musculus
557	31	63.3	431	2	Q66I16_XENTR	Q66i16 xenopus tro
558	31	63.3	433	2	Q74M16_NANEQ	Q74mi6 nanoarchaeu
559	31	63.3	434	2	Q8AVM2_XENLA	Q8avm2 xenopus lae
560	31	63.3	435	2	Q8I5V3_PLAF7	Q8i5v3 plasmodium
561	31	63.3	437	2	Q5X7E3_LEGPA	Q5x7e3 legionella
562	31	63.3	437	2	Q52XX0_LEGPH	Q52xx0 legionella
563	31	63.3	450	2	Q96MS1_HUMAN	Q96ms1 homo sapien
564	31	63.3	451	1	Y413_STRAS	Q86e1 streptococc
565	31	63.3	451	1	Y448_STRAS	Q86e1 streptococc
566	31	63.3	452	2	Q75IY1_ORYSA	Q75iy1 oryza sativ
567	31	63.3	452	2	Q56630_VIBCH	Q56630 vibrio chol
568	31	63.3	452	2	Q6P846_XENTR	Q6p846 xenopus tro
569	31	63.3	459	2	Q8EQ57_OCEIH	Q8eq57 oceanobacil
570	31	63.3	459	2	Q4SVZ8_TETNG	Q4svz8 tetraodon n
571	31	63.3	460	2	Q5AMT5_EMENI	Q5awt5 aspergillus
572	31	63.3	462	2	Q5AQ24_DICDI	Q5aq24 dictyosteli
573	31	63.3	464	2	Q4NT34_9DELT	Q4nt34 anaeromyxob
574	31	63.3	466	2	Q4NQ08_BACCE	Q4nq08 bacillus ce
575	31	63.3	474	2	Q4I5V1_GIBZE	Q4i5v1 gibberella
576	31	63.3	477	2	Q5P0X2_AZOSE	Q5p0x2 azocarcus sp
577	31	63.3	477	2	Q7NYV2_CHRVO	Q7nyv2 chromobacte
578	31	63.3	479	2	Q6BTT2_DBBHA	Q6btt2 debaryomyce
579	31	63.3	486	2	Q5JLGO_ORYSA	Q5jlgo oryza sativ
580	31	63.3	488	1	ARLY2_RHIME	Q92vm6 rhizobium m
581	31	63.3	488	1	U2AF2_CAEBR	P90727 caenorhabdi
582	31	63.3	496	2	Q5GCA1_RHOGE	Q5gca4 rhodocyclus
583	31	63.3	496	2	Q5WVS0_LEGPA	Q5wvs0 legionella
584	31	63.3	496	2	Q5X4D5_LEGPA	Q5x4d5 legionella
585	31	63.3	496	2	Q5ZUL6_LEGPH	Q5zul6 legionella
586	31	63.3	501	2	Q6GJD5_KLULA	Q6gjd5 klyveromyc
587	31	63.3	501	2	Q8G7L4_BIFLO	Q8g7l4 bifidobacte
588	31	63.3	501	2	Q896U3_CLOTE	Q896u3 clostridium
589	31	63.3	502	2	Q9VFS1_DROME	Q9vfs1 drosophila
590	31	63.3	507	2	Q9N5Z9_CABEL	Q9n5z9 caenorhabdi
591	31	63.3	514	2	Q9LJY6_ARATH	Q9ljy6 arabidopsis
592	31	63.3	515	2	Q4LX93_9BURK	Q4lx93 burkholderi
593	31	63.3	518	2	Q9YK9_XENLA	Q9yk9 xenopus lae
594	31	63.3	520	2	Q4LDN2_XENLA	Q4ldn2 xenopus lae
595	31	63.3	527	2	Q87G39_VIBPA	Q87g39 vibrio para
596	31	63.3	535	2	Q4Z6E2_PLABE	Q4z6e2 plasmodium
597	31	63.3	537	2	Q804W8_FUGRU	Q804w8 fugu rubrip
598	31	63.3	543	2	Q95JUS5_MACPA	Q95j55 macaca fasc
599	31	63.3	549	2	Q67M51_SYMTH	Q67m51 symbiobacte
600	31	63.3	557	2	Q54AN6_DICDI	Q54an6 dictyosteli
601	31	63.3	557	2	Q73766_XENLA	Q73766 xenopus lae
602	31	63.3	564	2	Q4I6V9_GIBZE	Q4i6v9 gibberella
603	31	63.3	578	2	Q9RGCS_MYCHY	Q9rgcs mycoplasma
604	31	63.3	590	2	Q9KVA7_VIBCH	Q9kva7 vibrio chol
605	31	63.3	593	2	Q8YPL5_ANASP	Q8ypl5 anabaena sp
606	31	63.3	597	2	Q6FXJ1_CANGA	Q6fxj1 candida gla
607	31	63.3	600	2	Q4SAX3_TETNG	Q4sax3 tetraodon n
608	31	63.3	601	2	Q4YB39_PLABE	Q4yb39 plasmodium
609	31	63.3	603	2	Q7RHJ1_PLAYO	Q7rhj1 plasmodium
610	31	63.3	617	2	Q6LKL8_PHOPR	Q6lkl8 photobacter
611	31	63.3	626	2	Q50J3P8_BRAKE	Q50j3p8 brachydanio
612	31	63.3	627	2	Q7T2M4_CARAU	Q7t2m4 carassius a
613	31	63.3	630	2	Q4I5G9_GIBZE	Q4i5g9 gibberella
614	31	63.3	632	1	ASNB_BACSU	P54420 bacillus su
615	31	63.3	636	2	Q5TW56_ANOGA	Q5tw56 anopheles g

689 31 63.3 1205 2 P92021 caenorhabdi
690 31 63.3 1210 2 Q8EQJ8_OCEIH
691 31 63.3 1215 2 Q8BI82_DROME
692 31 63.3 1257 2 Q55P94_CRYNE
693 31 63.3 1280 2 Q6FNG0_CANGA
694 31 63.3 1307 2 Q4QDX6_LBIMA
695 31 63.3 1317 2 Q7RH57_PLAYO
696 31 63.3 1322 2 Q54TW7_DICDI
697 31 63.3 1322 2 Q55D46_DICDI
698 31 63.3 1322 2 Q55GX9_DICDI
699 31 63.3 1335 2 Q55GX6_DICDI
700 31 63.3 1382 2 Q4SM61_TETNG
701 31 63.3 1423 2 Q5TJ22_HUMAN
702 31 63.3 1461 2 Q7RB11_PLAYO
703 31 63.3 1509 2 Q8S6P1_ORYSA
704 31 63.3 1534 2 Q4TB75_TETNG
705 31 63.3 1535 2 Q5CIU2_CRYHO
706 31 63.3 1692 2 Q8IJDE_PLAF7
707 31 63.3 2057 2 Q882M4_PSESM
708 31 63.3 2087 2 Q4S488_TETNG
709 31 63.3 2190 2 Q9VJMO_DROME
710 31 63.3 2223 1 CAC1E_DISOM
711 31 63.3 2314 2 Q7PQ11_ANOGA
712 31 63.3 2576 2 Q8JFV5_BRARE
713 31 61.2 35 2 Q8RIW2_CLODI
714 30 61.2 58 2 Q7MBJ5_VIBVY
715 30 61.2 72 2 Q9N3J5_CABEL
716 30 61.2 79 2 Q500F1_PSEBSY
717 30 61.2 87 2 Q05325_PSEFL
718 30 61.2 89 2 Q4LHK1_9BURK
719 30 61.2 90 2 Q86WH1_HUMAN
720 30 61.2 93 2 Q30682_RHOER
721 30 61.2 93 2 Q3WY27_LEGPL
722 30 61.2 93 2 Q5X7J9_LEGPA
723 30 61.2 93 2 Q5ZY31_LEGPH
724 30 61.2 98 2 Q8BL82_PSEPK
725 30 61.2 101 2 Q8ND09_HUMAN
726 30 61.2 110 2 P95874_SULSO
727 30 61.2 110 2 Q5C0D4_SCHJA
728 30 61.2 111 2 Q5BF85_ENENI
729 30 61.2 111 2 Q9WMG9_PHV
730 30 61.2 115 1 VGS2_SHV21
731 30 61.2 115 2 Q80B75_9GAMA
732 30 61.2 115 2 Q80BR9_9GAMA
733 30 61.2 115 2 Q80BR6_9GAMA
734 30 61.2 116 2 Q6MS32_MYCMS
735 30 61.2 118 1 RS13_YERPE
736 30 61.2 118 2 Q664U3_YERPS
737 30 61.2 119 2 Q8CX75_OCEIH
738 30 61.2 120 2 Q7NQ84_CHRVO
739 30 61.2 123 2 Q95Y08_CABEL
740 30 61.2 126 1 GRP_BOVIN
741 30 61.2 126 2 Q6AIH7_DESPS
742 30 61.2 129 2 Q5HML8_STAEO
743 30 61.2 131 2 Q61T90_CABBR
744 30 61.2 133 2 Q6DHG4_BRARE
745 30 61.2 134 1 GRP_SHEEP
746 30 61.2 134 2 Q94HX5_ORYSA
747 30 61.2 134 2 Q7XCR6_ORYSA
748 30 61.2 134 2 Q4FLU0_9RICK
749 30 61.2 134 2 Q6EMG6_9GEMI
750 30 61.2 134 2 Q9WR11_9GEMI
751 30 61.2 134 2 Q58WJ1_9GEMI
752 30 61.2 134 2 Q58WJ7_9GEMI
753 30 61.2 136 2 Q753T6_ASHGO
754 30 61.2 137 2 Q91058_HOPLIAS sp.
755 30 61.2 140 2 Q83028_DEIRA
756 30 61.2 144 2 Q8YJU7_ANASP
757 30 61.2 149 2 Q85698_STRLI
758 30 61.2 149 2 Q7AKR5_STRCO
759 30 61.2 149 2 Q92RB8_RHIME
760 30 61.2 150 1 RS13_CANNA
761 30 61.2 150 1 RS13_SCHPO

P92021 caenorhabdi
Q8eqj8 oceanobacil
Q5b1b2 drosophila
Q55p94 cryptococcu
Q6fng0 candida gla
Q4qdx6 leishmania
Q7rh57 plasmodium
Q54tw7 dictyosteli
Q55d46 dictyosteli
Q55gx9 dictyosteli
Q55gx6 dictyosteli
Q4sm61 tetraodon n
Q8t2j2 homo sapien
Q7rb11 plasmodium
Q8s6p1 oryza sativ
Q4tb75 tetraodon n
Q5ci22 cryptospori
Q8ijde plasmodium
Q882m4 pseudomonas
Q4s488 tetraodon n
Q9vjmo drosophila
P56699 discopyge o
Q7pql1 anopheles g
Q8jfv5 brachydanio
Q8riw2 clostridium
Q7mbj5 vibrio vuln
Q9n3j5 caenorhabdi
Q500f1 pseudomonas
Q05325 pseudomonas
Q4lhk1 burkholderi
Q86wh1 homo sapien
Q30682 rhodococcus
Q5wyz7 legionella
Q5zy31 legionella
Q88l82 pseudomonas
Q8nd09 homo sapien
P95874 sulfolobus
Q5c0d4 schistosoma
Q5bfe5 aspergillus
Q9wmg9 prospect hi
Q01050 saimirine
Q80bj5 saimirine
Q80br6 saimirine
Q80br9 saimirine
Q6ms32 mycoplasma
Q8zj90 versinia ps
Q664u3 versinia ps
Q8cx75 oceanobacil
Q7nq84 chromobacte
Q95y08 caenorhabdi
Q863c3 bos taurus
Q6aih7 desulfotale
Q5hml8 staphylococ
Q61f0 caenorhabdi
Q6dhg4 brachydanio
P47851 ovis aries
Q94hx5 oryza sativ
Q7xcr6 oryza sativ
Q4fluo candidatus
Q6emg6 east africa
Q9wr11 east africa
Q58wj1 east africa
Q58wj7 east africa
Q753t6 ashbya gos8
Q91058 hoplias sp.
Q83028 deinococcus
Q8yju7 anabaena sp
Q85698 streptomyce
Q7akr5 streptomyce
Q92rb8 rhizobium m
P33192 candida mal
P28189 schistosacch

762 30 61.2 151 2 Q6C242_YARLI
763 30 61.2 151 2 Q7VI02_HELHP
764 30 61.2 151 2 Q55NH9_BACLD
765 30 61.2 154 2 Q26394_METTH
766 30 61.2 154 2 Q7NJZ4_GLOVI
767 30 61.2 162 2 Q72CA7_DESVH
768 30 61.2 172 2 Q6ASD4_PROAC
769 30 61.2 172 2 Q91057_9TELE
770 30 61.2 175 2 Q4MRP1_BACCE
771 30 61.2 175 2 Q8HJV9_BACHK
772 30 61.2 175 2 Q81RN5_BACAN
773 30 61.2 175 2 Q91140_9TELE
774 30 61.2 178 2 Q4JA49_SULAC
775 30 61.2 178 2 Q91083_9TELE
776 30 61.2 181 2 Q90394_9TELE
777 30 61.2 181 2 Q91109_9TELE
778 30 61.2 188 2 Q5XG31_XENLA
779 30 61.2 190 2 Q5SPAI_BRARE
780 30 61.2 192 2 Q51CP4_ENTHI
781 30 61.2 194 2 Q5T5S0_HUMAN
782 30 61.2 194 2 Q9KPR4_VIBCH
783 30 61.2 194 2 Q4SS57_TETNG
784 30 61.2 199 2 Q9RMF1_ZYMMO
785 30 61.2 200 2 Q7M7V3_WOLSU
786 30 61.2 201 2 Q8FJG8_ECOL6
787 30 61.2 201 2 Q8XG55_SALTI
788 30 61.2 201 2 Q8RSJ1_MARMO
789 30 61.2 205 2 Q87GF6_VIBPA
790 30 61.2 207 2 Q73612_BACC1
791 30 61.2 207 2 Q813H0_BACCR
792 30 61.2 210 2 Q625G0_CABBR
793 30 61.2 210 2 Q9SYD8_ARATH
794 30 61.2 210 2 Q5PSE0_9CHLO
795 30 61.2 213 1 STARS_MOUSE
796 30 61.2 213 2 Q9CQY0_MOUSE
797 30 61.2 214 2 Q6FJU9_CANGA
798 30 61.2 216 2 Q4MPD0_BACCE
799 30 61.2 217 2 Q4X744_PLACH
800 30 61.2 218 1 Y395_PASMU
801 30 61.2 219 2 Q64CK5_9ARCH
802 30 61.2 219 2 Q5PSF0_9CHLO
803 30 61.2 220 2 Q50LX9_ENTHI
804 30 61.2 226 1 Y1442_YERPE
805 30 61.2 226 1 Y1460_YERPS
806 30 61.2 229 2 P94406_BACSU
807 30 61.2 230 2 Q6P2D9_HUMAN
808 30 61.2 230 2 Q72QV2_LEPIC
809 30 61.2 231 2 Q4XD89_PLACH
810 30 61.2 232 2 Q99P51_MOUSE
811 30 61.2 238 2 Q86WH2_HUMAN
812 30 61.2 238 2 Q8NPL3_CORGL
813 30 61.2 241 1 PHOU_ECOLI
814 30 61.2 241 1 PHOU_SHIFL
815 30 61.2 241 1 PHOU_SHIFL
816 30 61.2 241 2 Q57HY9_SALCH
817 30 61.2 241 2 Q5PKV0_SALPA
818 30 61.2 241 2 Q7CPE7_SALTY
819 30 61.2 241 2 Q8XF10_SALTI
820 30 61.2 241 2 Q8XB06_ECOS7
821 30 61.2 242 2 Q8F4Z6_LEPIN
822 30 61.2 243 2 Q4HTN6_CAMUP
823 30 61.2 246 2 Q8S7R9_ORYSA
824 30 61.2 247 2 Q4N8W6_THEPA
825 30 61.2 247 2 Q8FST8_ECOL6
826 30 61.2 249 2 Q4UER8_THEAN
827 30 61.2 249 2 Q9K710_BACHD
828 30 61.2 251 2 Q92V08_RHIME
829 30 61.2 251 2 Q5PSE1_9CHLO
830 30 61.2 251 2 Q5PSF4_9CHLO
831 30 61.2 252 2 Q5PSE8_9CHLO
832 30 61.2 253 2 Q50P50_ENTHI
833 30 61.2 256 1 Y022_THEAC
834 30 61.2 256 2 Q5JEG3_PYRKO

Q6c242 varrowia li
Q7vi02 helicobacte
Q55nh9 bacillus li
Q26394 methanobact
Q7njz4 gloeobacter
Q72ca7 desulfovibr
Q6as44 propionibac
Q91057 bacillus sp
Q4mrfl bacillus ce
Q8hvj9 bacillus th
Q81rn5 bacillus an
Q91140 nannobrycon
Q4ja49 sulfolobus
Q91083 leporinus s
Q90394 chalcus sp
Q91109 metynnis sp
Q5xg31 xenopus lae
Q5spai brachydanio
Q51cp4 entamoeba h
Q5t5s0 homo sapien
Q9kpr4 vibrio chol
Q4ss57 tetraodon n
Q9rmf1 zymomonas m
Q7m7v3 wolnelia s
Q8fjg8 escherichia s
Q8xg55 salmonella
Q8rsj1 marmota mon
Q87gf6 vibrio para
Q73612 bacillus ce
Q813h0 bacillus ce
Q625g0 caenorhabdi
Q9syd8 arabidopsis
Q5pse0 halimeda in
Q9cpq7 mus musculu
Q6fju9 candida gla
Q4mpd0 bacillus ce
Q4x744 plasmodium
Q6cm22 pasteurella
Q64ck5 uncultured
Q5psf0 halimeda in
Q50lx9 entamoeba h
Q8zg70 versinia pe
Q66c3 versinia ps
Q99p51 mus musculu
Q86wh2 homo sapien
Q8npl3 corynebacte
P0a9k7 escherichia
Q32498 enterobacte
P0a9k8 shigella fl
Q57hy9 salmonella
Q5pkv0 salmonella
Q7cpe7 salmonella
Q8xf10 salmonella
Q8xb06 escherichia
Q8f4z6 leptospira
Q4htn6 campylobact
Q8s7r9 oryza sativ
Q4n8w6 theileria p
Q8fb8 escherichia
Q9uer8 theileria a
Q9k710 bacillus ha
Q92v08 rhizobium m
Q5pse1 halimeda in
Q5psf4 halimeda in
Q5pse8 halimeda in
Q50p50 entamoeba h
Q9hm47 thermoplasma
Q5jeg3 pyrococcus

835	30	61.2	256	2	Q5YD3_PVRAB	Q9uyd3 pyrococcus	908	61.2	354	2	Q95Y07_CAEEL	Q95y07 caenorhabdi
836	30	61.2	256	2	Q58212_PVRHO	Q58212 pyrococcus	909	61.2	357	2	Q6CBE2_YARLI	Q6cb2 yarrowia li
837	30	61.2	257	2	Q5L0M4_GROKA	Q5l0m4 geobacillus	910	61.2	359	2	Q60759_HUMAN	Q60759 homo sapien
838	30	61.2	257	2	Q8U3K7_PFRFU	Q8u3k7 pyrococcus	911	61.2	359	2	Q8NE32_HUMAN	Q8ne32 homo sapien
839	30	61.2	257	2	Q6HIW4_BACHK	Q6hiw4 bacillus th	912	61.2	361	2	Q7S261_BRARE	Q7s261 brachydanio
840	30	61.2	258	2	Q5PSF1_9CHLO	Q5psf1 halimeda in	913	61.2	362	2	Q9CTC6_CAEEL	Q9ctc6 caenorhabdi
841	30	61.2	258	2	Q5PSF6_9CHLO	Q5psf6 halimeda in	914	61.2	363	2	Q5KFS1_CRYNE	Q5kfs1 cryptococcu
842	30	61.2	258	2	Q6MQJ6_BDEBA	Q6mqj6 bdellovibri	915	61.2	364	2	Q50UT8_ENTHI	Q50ut8 entameoba h
843	30	61.2	260	2	Q5PSB9_9CHLO	Q5psb9 halimeda si	916	61.2	364	2	Q8VI91_RAT	Q8vi91 rattus norv
844	30	61.2	260	2	Q5PSB4_9CHLO	Q5psb4 halimeda in	917	61.2	365	2	Q9LHR3_ARATH	Q9lhr3 arabidopsis
845	30	61.2	261	2	Q4YSN2_PLACH	Q4ysn2 plasmodium	918	61.2	367	2	Q9KLV7_VIBCH	Q9klv7 vibrio chol
846	30	61.2	262	1	VIBA_VIBCH	Q56632 vibrio chol	919	61.2	370	2	Q6LIU3_PHOPE	Q6liu3 photobacter
847	30	61.2	262	2	Q5PSE5_9CHLO	Q5pse5 halimeda in	920	61.2	371	2	Q9VKE8_DROME	Q9vke8 drosophila
848	30	61.2	263	2	Q5PSE6_9CHLO	Q5pse6 halimeda in	921	61.2	373	2	Q6CPK4_KLULA	Q6cpk4 kluyveromyc
849	30	61.2	263	2	Q5PSE7_9CHLO	Q5pse7 halimeda in	922	61.2	375	2	Q9CKX1_MOUSE	Q9ckx1 mus musculu
850	30	61.2	263	2	Q5PSE9_9CHLO	Q5pse9 halimeda in	923	61.2	379	2	Q738T6_BACCI	Q738t6 bacillus ce
851	30	61.2	267	2	Q9CT54_MOUSE	Q9ct54 mus musculu	924	61.2	379	2	Q5PPN7_RAT	Q5ppn7 rattus norv
852	30	61.2	273	2	Q83507_TREPA	Q83507 treponema p	925	61.2	380	2	Q96ER9_HUMAN	Q96er9 homo sapien
853	30	61.2	275	2	Q6BPP6_PARTE	Q6bpf6 paramecium	926	61.2	382	1	METK_LEGPA	Q8x3n0 legionella
854	30	61.2	275	2	Q5PSB8_9CHLO	Q5psb8 halimeda si	927	61.2	382	1	METK_LEGPH	Q5ztv6 legionella
855	30	61.2	275	2	Q5PSC1_9CHLO	Q5psc1 halimeda si	928	61.2	382	1	METK_LEGPL	Q5wv18 legionella
856	30	61.2	275	2	Q5PSC2_9CHLO	Q5psc2 halimeda si	929	61.2	382	1	Q4N977_THEPA	Q4n977 theileria p
857	30	61.2	275	2	Q5PSC3_9CHLO	Q5psc3 halimeda si	930	61.2	382	2	Q64SE2_BACPR	Q64se2 bacteroides
858	30	61.2	275	2	Q5PSC5_9CHLO	Q5psc5 halimeda si	931	61.2	382	2	Q9A7B7_CAUCR	Q9a7b7 caulobacter
859	30	61.2	275	2	Q5PSC7_9CHLO	Q5psc7 halimeda si	932	61.2	382	2	Q5L3D7_GEOCA	Q5l3d7 geobacillus
860	30	61.2	279	2	Q4R6X3_MACFA	Q4r6x3 macaca fasc	933	61.2	385	2	Q98K34_RHILO	Q98k34 rhizobium l
861	30	61.2	284	2	Q2VNS5_ARATH	Q2vns5 arabidopsis	934	61.2	385	2	Q55ZU0_CRYNE	Q55zu0 cryptococcu
862	30	61.2	286	2	Q6ALV3_DESPS	Q6alv3 desulfotale	935	61.2	387	2	P95884_SULSO	P95884 sulfobolus
863	30	61.2	286	2	Q6N863_RHOPA	Q6n863 rhodopseudo	936	61.2	387	2	Q7QC53_ANOGA	Q7qc53 anopheles g
864	30	61.2	288	2	Q7TSY5_9PICO	Q7tsy5 human rhino	937	61.2	388	2	Q9IDT2_VIRU	Q9idt2 rice tungro
865	30	61.2	292	2	Q8XJ66_CLOPE	Q8xj66 clostridium	938	61.2	390	2	Q96KY9_HUMAN	Q96ky9 homo sapien
866	30	61.2	295	2	Q7CME8_BACAN	Q7cme8 bacillus an	939	61.2	390	2	Q5H825_STRSU	Q5h825 streptococc
867	30	61.2	295	2	Q4MXK0_BACCE	Q4mxk0 bacillus ce	940	61.2	390	2	Q4NMV4_9DELJ	Q4nmv4 anaeromyxob
868	30	61.2	295	2	Q9X364_BACAN	Q9x364 bacillus a	941	61.2	392	2	Q7QU72_GIALA	Q7qu72 giardia lam
869	30	61.2	297	2	Q4UGH8_THEAN	Q4ugh8 theileria a	942	61.2	392	2	Q9HP92_HALSA	Q9hp92 halobacteri
870	30	61.2	298	1	ERA_XANCP	Q8pb51 xanthomonas	943	61.2	397	2	Q4UHL9_THEAN	Q4uhl9 theileria a
871	30	61.2	298	2	Q7S3S2_NEUCR	Q7s3s2 neurospora	944	61.2	401	2	Q5T5S1_HUMAN	Q5t5s1 homo sapien
872	30	61.2	298	2	Q4USP8_XANCP	Q4usf8 xanthomonas	945	61.2	403	2	Q8ECA4_SHEON	Q8eca4 shewanella
873	30	61.2	299	1	ERA_XANAC	Q8pmu9 xanthomonas	946	61.2	404	2	Q5E9Q3_BOVIN	Q5e9q3 bos taurus
874	30	61.2	299	2	Q5H1R1_XANOR	Q5h1r1 xanthomonas	947	61.2	404	2	Q5E9W7_BOVIN	Q5e9w7 bos taurus
875	30	61.2	301	2	Q5CE52_CRYHO	Q5ce52 cryptospori	948	61.2	404	2	Q5EA69_BOVIN	Q5ea69 bos taurus
876	30	61.2	302	2	Q9HA01_HUMAN	Q9ha01 homo sapien	949	61.2	404	2	Q6NDN6_RHOPA	Q6ndn6 rhodopseudo
877	30	61.2	303	1	CBPB_ASTFL	P04069 astacae flu	950	61.2	406	1	DAC_STRSR	P15555 streptomyce
878	30	61.2	305	2	Q5ACE1_DICDI	Q5acel dictyosteli	951	61.2	406	2	Q4QQL1_MOUSE	Q4qql1 mus musculu
879	30	61.2	305	2	Q7GYR2_CAEEL	Q7gyr2 caenorhabdi	952	61.2	409	2	Q61103_TYRCR	Q61103 trypanosoma
880	30	61.2	306	2	Q61LR8_CAEER	Q61lr8 caenorhabdi	953	61.2	415	2	Q4S8M8_TPTNG	Q4s8m8 tetraodon n
881	30	61.2	309	2	Q4LLB7_9BURK	Q4llb7 burkholderi	954	61.2	417	2	Q8YPR6_ONYPE	Q8ypr6 onion yello
882	30	61.2	313	1	MDH_BACTN	Q8abw0 bacteroides	955	61.2	418	2	Q8TK44_METAC	Q8tk44 methanosarc
883	30	61.2	313	2	Q7TRP1_MOUSE	Q7trp1 mus musculu	956	61.2	421	2	Q70J68_STRGR	Q70j68 streptomyce
884	30	61.2	314	2	Q6JVA7_ANTPO	Q6jva7 antheraea p	957	61.2	422	2	Q7R4K4_GIALA	Q7r4k4 giardia lam
885	30	61.2	315	2	Q5JG61_PYRKO	Q5jg61 pyrococcus	958	61.2	422	2	Q6ZUX0_HUMAN	Q6zux0 homo sapien
886	30	61.2	316	2	Q50T39_ENTHI	Q50t39 entamoeba h	959	61.2	423	2	Q5CZ27_CRYPV	Q5cz27 cryptospori
887	30	61.2	318	2	Q7N1K9_GLOBO	Q7nik9 gloeobacter	960	61.2	423	2	Q5CMB9_PORGI	Q5cmb9 porphyromon
888	30	61.2	319	2	Q4H5T7_9DEIO	Q4h5t7 deinococcus	961	61.2	427	2	Q7MTB1_PORGI	Q7mtb1 porphyromon
889	30	61.2	321	1	RLAO_CHERU	P29764 chenopodium	962	61.2	430	1	SYH_LACLA	Q9ce78 lactococcus
890	30	61.2	323	2	Q6BG45_PARTE	Q6bg45 paramecium	963	61.2	432	2	Q6G8M2_STAAS	Q6g8m2 staphylococ
891	30	61.2	324	1	MURB_RHIME	Q924m1 rhizobium m	964	61.2	432	2	Q6GG02_STAAR	Q6gg02 staphylococ
892	30	61.2	324	2	Q5B8C7_EMENI	Q5bbc7 aspergillus	965	61.2	432	2	Q8HF69_STAAC	Q8hf69 staphylococ
893	30	61.2	325	2	Q50N53_ENTHI	Q50n53 entamoeba h	966	61.2	432	2	Q7A0N2_STAAN	Q7a0n2 staphylococ
894	30	61.2	326	2	Q7Q7V1_ANOGA	Q7q7v1 anopheles g	967	61.2	432	2	Q7A554_STAAN	Q7a554 staphylococ
895	30	61.2	326	2	Q50FQ4_ENTHI	Q50fq4 entamoeba h	968	61.2	432	2	Q99TF8_STAAM	Q99tf8 staphylococ
896	30	61.2	327	1	TRXB_PNEJI	Q8j0u0 pneumocysti	969	61.2	433	2	Q6FTS3_CANGA	Q6fts3 candida gla
897	30	61.2	331	2	Q9M4Z5_SPIOL	Q9m4z5 spinacia ol	970	61.2	433	2	Q50NDA_ENTHI	Q50nda entamoeba h
898	30	61.2	331	2	Q6DB83_ERWCT	Q6db83 erwinia car	971	61.2	434	2	Q9LH92_ARATH	Q9lhn2 arabidopsis
899	30	61.2	334	2	Q6N255_RHOPA	Q6n255 rhodopseudo	972	61.2	440	2	Q9LH93_ARATH	Q9lhn3 arabidopsis
900	30	61.2	335	2	Q22666_CAEEL	Q22666 caenorhabdi	973	61.2	440	2	Q74BH4_GEOSL	Q74bh4 geobacter s
901	30	61.2	335	2	Q7UNQ7_RHOBA	Q7unq7 rhodopirell	974	61.2	445	1	Y1571_PASMU	Q9mbg7 arabidopsis
902	30	61.2	335	2	Q4RSF1_TETNG	Q4rsf1 tetraodon n	975	61.2	446	1	Y1571_PASMU	Q9stks drosophila
903	30	61.2	336	2	Q5PHY0_LACAC	Q5phy0 lactobacill	976	61.2	448	2	Q9XWK5_DROME	Q9xwk5 caenorhabdi
904	30	61.2	339	2	Q8TF18_HUMAN	Q8tf18 homo sapien	977	61.2	448	2	Q9XVW6_CAEEL	Q9xvw6 caenorhabdi
905	30	61.2	340	2	Q26744_9TRYP	Q26744 trypanosoma	978	61.2	455	2	Q4KL34_STRPN	Q4kl34 streptococ
906	30	61.2	348	1	LIPA_CORGL	Q8nuj0 corynebacte	979	61.2	455	2	Q8KC87_CHLTE	Q8kc87 chlorobium
907	30	61.2	349	1	LIPA_CORDI	P61194 corynebacte	980	61.2	459	2	Q8NIB5_TALEM	Q8nib5 talaromyces

981 30 61.2 459 2 Q6ST59 human rhino
 982 30 61.2 462 2 Q5QNU2 mus musculus
 983 30 61.2 463 1 VCL_VICFA
 984 30 61.2 466 2 Q9V2D8 Pyrodictus
 985 30 61.2 466 2 Q4QDW8 leishmania
 986 30 61.2 467 2 Q61DG1 caenorhabditis
 987 30 61.2 469 2 Q6UNX5 nicotiana b
 988 30 61.2 470 2 Q9X1N3 theca
 989 30 61.2 473 2 Q9JN22 bradyrhizob
 990 30 61.2 481 2 Q9UWX2 sulfolobus
 991 30 61.2 481 2 Q6BP54 debaryomyces
 992 30 61.2 483 1 UXAB_YERPE
 993 30 61.2 483 2 Q665N9 versinia ps
 994 30 61.2 489 2 Q9HMJ6 halobacterium
 995 30 61.2 490 2 Q9W6G5 brachydanio
 996 30 61.2 496 2 Q9HIK3 theca
 997 30 61.2 497 2 Q61101 trypanosoma
 998 30 61.2 501 2 Q94450 schistosoma
 999 30 61.2 505 2 Q8IA46 spodoptera
 1000 30 61.2 505 2 Q8IU00 caenorhabditis

ALIGNMENTS

RESULT 1
 Q4S6F3_TETNG
 ID Q4S6F3_TETNG PRELIMINARY; PRT; 458 AA.
 AC Q4S6F3;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome 10 SCAF14728, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00023316001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castellini V., Katinka M., Vacherie B.,
 RA Bismont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope, Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAAB01014728; CAG03779.1; -; Genomic_DNA.
 FT NON_TPR 458 458
 SQ SEQUENCE 458 AA; 51958 MW; 088CDE5D6E5F977 CRC64;

Query Match 83.7%; Score 41; DB 2; Length 458;
 Best Local Similarity 70.0%; Pred. NO. 34;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 ||:||||:|
 Db 307 ESIREQIKKY 316
 RESULT 2
 Q9ZJN3_HELPJ
 ID Q9ZJN3_HELPJ PRELIMINARY; PRT; 1164 AA.
 AC Q9ZJN3;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative.
 GN OrderedLocusNames=JHP1272;
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 DR EMBL; AE001550; AAD06844.1; -; Genomic_DNA.
 DR PIR; G71827; G71827.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008170; P:N-methyltransferase activity; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR InterPro; IPR002296; N12N6_mtfase.
 DR PRINTS; PR00507; N12N6MTFRASE.
 KW Complete proteome.
 SQ SEQUENCE 1164 AA; 134378 MW; 5EC6286EA439873E CRC64;

Query Match 81.6%; Score 40; DB 2; Length 1164;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
 ||:||||:|
 Db 740 ESLEQVILKY 749

RESULT 3
 Q4SFR7_TETNG
 ID Q4SFR7_TETNG PRELIMINARY; PRT; 1064 AA.
 AC Q4SFR7;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome 7 SCAF14601, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00018981001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castellini V., Katinka M., Vacherie B.,
 RA Bismont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAG01014601; 1064
DR NON TER 1064 AA; 1064
SQ SEQUENCE 1064 AA; 116866 MW; AF3106816A29D570 CRC64;

Query Match 79.6%; Score 39; DB 2; Length 1064;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|||||
DB 986 ESIREHVIKY 995

RESULT 4
ERA MYCPN STANDARD; PRT; 291 AA.
AC P75110;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE GTP-binding protein era homolog.
GN Name=era; Synonyms=spg; OrderedLocusNames=MPN568; ORFNames=MP274;
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=12104;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;
RA Himmelreich R., Hilbert H., Plegens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- FUNCTION: Binds both GDP and GTP, has an intrinsic GTPase activity
and is essential for cell growth (By similarity).
CC -!- SIMILARITY: Belongs to the era/trme GTP-binding protein family.
CC Era subfamily.
CC -!- SIMILARITY: Contains 1 KH type-2 domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; ABG000027; ABG95922.1; -; Genomic_DNA.
DR FIR; S73600; S73600.
DR HAMAP; MF_00367; -; 1.
DR InterPro; IPR005662; Era GTP bd.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR004044; KH_TYPE_2.
DR InterPro; IPR002917; MMR_HSR1_GTP_bd.
DR Pfam; PF07650; KH_2; 1.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PR00326; GTP1OGB.
DR TIGRFAMs; TIGR00436; era; 1.
DR PROSITE; PS50823; KH_TYPE_2; 1.

KW Complete proteome; GTP-binding; Nucleotide-binding; RNA-binding.
FT DOMAIN 201 279 KH type-2.
FT NP_BIND 10 17 GTP (Potential).
FT NP_BIND 58 62 GTP (Potential).
FT NP_BIND 122 125 GTP (Potential).
SQ SEQUENCE 291 AA; 33529 MW; 2E81B0406A9106E3 CRC64;

Query Match 77.6%; Score 38; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 79;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|||||
DB 191 EALREQUIKY 200

RESULT 5
Q52A99 MAGGR
ID Q52A99_MAGGR PRELIMINARY; PRT; 592 AA.
AC Q52A99;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG02135.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguelavskiy L. N.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseelis M., Karlsson E.,
RA Kellis C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokytsang T., Lokytsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihailescu A., Mihova T., Mikkelsen T., Mieng V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Steason K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wanchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [12]

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AACU0100408; EAA54150.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 592 AA; 68710 MW; 68F2DEBFBBC9E9EC CRC64;

Query Match 77.6%; Score 38; DB 2; Length 592;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 540 EFVREIRIMKY 549

RESULT 6
Q5CP31_CRYHO
ID Q5CP31_CRYHO PRELIMINARY; PRT; 593 AA.
AC Q5CP31;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE IP953.
GN ORFNames=Chro.70176;
OS Cryptosporidium hominis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=237895;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TUS02;
RA Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G.,
RA Putu D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H.,
RA Bankier A.T., Peterson D.L., Abrahamson M.S., Kapur V., Tzipori S.,
RA Buck G.A.;
RT "The genome of Cryptosporidium hominis."
RL Nature 431:1107-1112(2004).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophyl-tRNA(Trp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC EMBL; AAE0100003; EAL38376.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:Ligase activity; IEA.
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
DR GO; GO:0006436; P:protein biosynthesis; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
KW Protein biosynthesis.
SQ SEQUENCE 593 AA; 67816 MW; E2FECA7A13470D8E CRC64;

Query Match 77.6%; Score 38; DB 2; Length 593;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
```

```
Db 483 ESIRNKIMKY 492

RESULT 7
Q5CYP8_CRYPV
ID Q5CYP8_CRYPV PRELIMINARY; PRT; 593 AA.
AC Q5CYP8;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Tryptophanyl-tRNA synthetase.
GN ORFNames=cgd7_1490;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RC PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahamte J.E., Zhu G.,
RA Lancot C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
RT parvum."
RL Science 304:441-445(2004).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophyl-tRNA(Trp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC EMBL; AAE0100001; EAK90203.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:Ligase activity; IEA.
DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
DR GO; GO:0006436; P:protein biosynthesis; IEA.
DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW ATP-binding; Aminoacyl-tRNA synthetase; Ligase; Nucleotide-binding;
KW Protein biosynthesis.
SQ SEQUENCE 593 AA; 67886 MW; ECE9587B2ECF6ECC CRC64;

Query Match 77.6%; Score 38; DB 2; Length 593;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 483 ESIRNKIMKY 492

RESULT 8
Q6JIZ0_MOUSE
ID Q6JIZ0_MOUSE PRELIMINARY; PRT; 2448 AA.
AC Q6JIZ0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyl-CoA carboxylase 2.
GN Name=Acacb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Heart;
```

RA Mao J., Wakil S.J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY451394; AAS31686.1; -; mRNA.
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR011761; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR005482; Biotin carb C.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR000022; Carboxyl trans.
DR InterPro; IPR011763; COA CT C.
DR InterPro; IPR011762; COA CT N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; Cphp synth_L_D2.
DR InterPro; IPR002114; HPr Serp S.
DR Pfam; PF02785; Biotin carb C; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF01039; Carboxyl trans; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS00975; ATP GRASP; 1.
DR PROSITE; PS00979; BC; 1.
DR PROSITE; PS00968; BIOTINYL LIPOYL; 1.
DR PROSITE; PS00989; COA CT_CTER; 1.
DR PROSITE; PS00980; COA CT_NTER; 1.
DR PROSITE; PS00866; CPASE 1; UNKNOWN 1.
DR PROSITE; PS00867; CPASE 2; UNKNOWN 1.
DR PROSITE; PS00589; PTS HPr SPR; UNKNOWN 1.
SQ SEQUENCE 2448 AA; 275666 MW; 0B8649F5D2CF1C8A CRC64;

Query Match 77.6%; Score 38; DB 2; Length 2448;
Best Local Similarity 70.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESRVQVMKY 10
|||:|:|:
DB 1599 ESRVDMVMRY 1608

RESULT 9
Q9LJZ1 ARATH PRELIMINARY; PRT; 394 AA.
AC Q9LJZ1;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Similarity to cytochrome P450.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosids; eutrosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=20363099; PubMed=10907853;
RX Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221 (2000).
[2]
RN NUCLEOTIDE SEQUENCE.
RP Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DBSJ databases.
RL -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC EMBL; AP000383; BAB01867.1; -; Genomic DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.

DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 394 AA; 44658 MW; 5A3AAD918246BBA4 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESRVQVMKY 10
|||:|:|:
DB 316 EEIREQALKY 325

RESULT 10
IF5 MOUSE
ID IF5 MOUSE STANDARD; PRT; 429 AA.
AC P59325;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN Name=EIF5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=C57BL/6, and FVB/N; TISSUE=Colon, and Eye;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smaluk D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=20363099; PubMed=10907853;
RX Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221 (2000).
[2]
RN NUCLEOTIDE SEQUENCE.
RP Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DBSJ databases.
RL -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC EMBL; AP000383; BAB01867.1; -; Genomic DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.

```

DR MGI:95309; Eif5.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
DR GTP-binding; Initiation factor; Nucleotide-binding; Phosphorylation;
KW Protein biosynthesis.
FT NP_BIND 27 34 GTP (Potential).
FT COMPBIAS 194 200 Asp/Glu-rich (highly acidic).
FT COMPBIAS 382 400 Asp/Glu-rich (highly acidic).
FT COMPBIAS 421 427 Asp-rich (acidic).
FT MOD_RES 387 387 Phosphoserine (By similarity).
FT MOD_RES 388 388 Phosphoserine (By similarity).
SQ SEQUENCE 429 AA; 48968 MW; C77D5DFB85C1DEF CRC64;

Query Match 75.5%; Score 37; DB 1; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
|:|:|:|:|

RESULT 11
ID RAT STANDARD; PRT; 429 AA.
AC Q07205;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN Name=Eif5;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 131-150; 215-242 AND
RP 333-353.
RC TISSUE=Liver;
RX MEDLINE=93219424; PubMed=8464924;
RA Das K., Chevesich J., Maitra U.;
RT "Molecular cloning and expression of cDNA for mammalian translation
initiation factor 5";
RL Proc. Natl. Acad. Sci. U.S.A. 90:3058-3062(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PHOSPHORYLATION SITES SER-387 AND SER-388.
RX MEDLINE=21851300; PubMed=11861906; DOI=10.1093/nar/30.5.1154;
RA Majumdar R., Bandyopadhyay A., Deng H., Maitra U.;
RT "Phosphorylation of mammalian translation initiation factor 5 (eIF5)
in vitro and in vivo.";
RL Nucleic Acids Res. 30:1154-1162(2002).
CC -!- FUNCTION: Catalyzes the hydrolysis of GTP bound to the 40S
ribosomal initiation complex (40S.mRNA.Met-tRNA[E].eIF-2.GTP) with
the subsequent joining of a 60S ribosomal subunit resulting in the
release of eIF-2 and the guanine nucleotide. The subsequent
joining of a 60S ribosomal subunit results in the formation of a
functional 80S initiation complex (80S.mRNA.Met-tRNA[E]).
CC -!- SUBUNIT: Monomer; but some authors state that this protein may not
be a monomer.
CC -!- SIMILARITY: Belongs to the eIF-2-beta/eIF-5 family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; L11651; AAA41112.1; -; mRNA.
DR EMBL; BC062398; AAH62398.1; -; mRNA.
DR PIR; A47305; A47305.
DR Ensembl; ENSRNOC00000010218; Rattus norvegicus.
DR GDI; 619861; Eif5.
DR GO; GO:0003743; F:translation initiation factor activity; IDA.
DR GO; GO:0006412; P:protein biosynthesis; TAS.
DR InterPro; IPR002735; eIF5_eIF2B.
DR InterPro; IPR003307; eIF5C.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
KW Direct protein sequencing; GTP-binding; Initiation factor;
KW Nucleotide-binding; Phosphorylation; Protein biosynthesis.
FT NP_BIND 27 34 GTP (Potential).
FT COMPBIAS 194 200 Asp/Glu-rich (highly acidic).
FT COMPBIAS 382 400 Asp/Glu-rich (highly acidic).
FT COMPBIAS 421 427 Asp-rich (acidic).
FT MOD_RES 387 387 Phosphoserine.
FT MOD_RES 388 388 Phosphoserine.
SQ SEQUENCE 429 AA; 48954 MW; B1A62E30936908EE CRC64;

Query Match 75.5%; Score 37; DB 1; Length 429;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
Db 285 EKIREQIKKY 294
|:|:|:|:|

RESULT 12
OSZIEO CHICK
ID OSZIEO_CHICK PRELIMINARY; PRT; 430 AA.
AC Q5ZIE0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=NCU0604_27j21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatovski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
gene function analysis.";
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ720844; CAG32503.1; -; mRNA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006446; P:regulation of translational initiation; IEA.
DR InterPro; IPR003307; eIF5C.
DR InterPro; IPR002735; eIF5_eIF2B.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR ProDom; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
DE Hypothetical protein.
SQ SEQUENCE 430 AA; 49078 MW; 6105AC7230CBA878 CRC64;

Query Match 75.5%; Score 37; DB 2; Length 430;

```


Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ESRVQVWKY 10
Db 285 EKIREQIRK 294

RESULT 13
ID IF5 HUMAN STANDARD; PRT; 431 AA.
AC P55010; Q9H5N2; Q9UG48;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 46, Last annotation update)
DE Eukaryotic translation initiation factor 5 (eIF-5).
GN Name=EIF5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96279275; PubMed=8663286; DOI=10.1074/jbc.271.28.16934;
RA Si K., Das K., Maitra U.;
RT "Characterization of multiple mRNAs that encode mammalian translation
initiation factor 5 (eIF-5).";
RL J. Biol. Chem. 271:16934-16938(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wienann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgor W., Boecker M., Blocker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koeher K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwawaki T., Watanabe M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togashi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Koniyama M., Tashiro H., Tanigami A., Fujizawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiraoka M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobayashi N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani K., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human

CDNAS.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- FUNCTION: Catalyzes the hydrolysis of GTP bound to the 40S
ribosomal initiation complex (40S.mRNA.Met-tRNA[F].eIF-2.GTP) with
the subsequent joining of a 60S ribosomal subunit resulting in the
release of eIF-2 and the guanine nucleotide. The subsequent
joining of a 60S ribosomal subunit results in the formation of a
functional 80S initiation complex (80S.mRNA.Met-tRNA[F]).
CC -!- INTERACTION:
CC -!- OIA602:EIF5; NbExp=1; IntAct=EBI-286450, EBI-286439;
CC -!- SIMILARITY: Belongs to the eIF-2-beta/eIF-5 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; U49436; AAC50572.1; -; mRNA.
CC EMBL; AL080102; CAB45711.1; -; mRNA.
CC EMBL; AK026933; BAB15593.1; -; mRNA.
CC EMBL; BC007728; AA077728.1; -; mRNA.
CC EMBL; BC032866; AA032866.1; -; mRNA.
CC PIR; T12450; T12450.
CC IntAct; P55010; -;
CC Ensembl; ENSG00000100664; Homo sapiens.
CC HGNC; HGNC:3299; EIF5.
CC H-InvDB; HIX0011999; -;
CC Reactome; P55010; -;
CC MIM; 601710; -;
CC GO; GO:0005829; C:cytosol; NAS.
CC GO; GO:0003924; F:GTPase activity; TAS.
CC GO; GO:0003743; P:translation initiation factor activity; NAS.
CC GO; GO:0006446; P:regulation of translational initiation; TAS.
CC InterPro; IPR002735; eIF5_eIF2B.
CC InterPro; IPR003307; eIF5C.
CC Pfam; PF01873; eIF-5_eIF-2B; 1.
CC Pfam; PF02020; W2; 1.
CC ProDom; PD004078; eIF5_eIF2B; 1.
CC SMART; SM00653; eIF2B_5; 1.
CC SMART; SM00515; eIF5C; 1.
CC GTP-binding; Initiation factor; Nucleotide-binding; Phosphorylation;
KW Protein biosynthesis.
FT NP_BTND 27 34 GTP (Potential).
FT COMEPIAS 196 202 Asp/Glu-rich (highly acidic).
FT COMEPIAS 384 402 Asp/Glu-rich (highly acidic).
FT COMEPIAS 423 429 Asp-rich (acidic).
FT MOD_RES 389 389 Phosphoserine (By similarity).
FT MOD_RES 390 390 Phosphoserine (By similarity).
FT

```
FT CONFLICT      60      60      E -> G (in Ref. 21).
FT CONFLICT     203     203      W -> S (in Ref. 1).
FT CONFLICT     295     295      K -> E (in Ref. 2).
FT CONFLICT     311     311      Q -> K (in Ref. 4; AAH32866).
SQ SEQUENCE     431 AA; 49223 MW; C6CCC3A255F9B9BC CRC64;

Query Match      75.5%; Score 37; DB 1; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
   |:|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 14
ID Q53XB3 HUMAN PRELIMINARY; PRT; 431 AA.
AC Q53XB3
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein DKFZp686K0727.
GN Name=DKFZp686K0727;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Human testis;
RA Anorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Meves H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX537367; CAD97610.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 49222 MW; C6CCC3A255F9B9BC CRC64;

Query Match      75.5%; Score 37; DB 2; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
   |:|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 15
ID Q5R4L0 PONPY PRELIMINARY; PRT; 431 AA.
AC Q5R4L0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp459F153.
GN Name=DKFZp459F153;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cortex;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Meves H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR861236; CAH93306.1; -; mRNA.
DR GO; GO:0003743; P:translation initiation factor activity; IEA.
DR GO; GO:0006446; P:regulation of translational initiation; IEA.
DR InterPro; IPR003307; eIF5C.
```

```
DR InterPro; IPR002735; eIF5_eIF2B.
DR Pfam; PF01873; eIF-5_eIF-2B; 1.
DR Pfam; PF02020; W2; 1.
DR PRODOM; PD004078; eIF5_eIF2B; 1.
DR SMART; SM00653; eIF2B_5; 1.
DR SMART; SM00515; eIF5C; 1.
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 49211 MW; 9789C6C01D7638CC CRC64;

Query Match      75.5%; Score 37; DB 2; Length 431;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
   |:|:|:|:|
Db 287 EKIREQIKKY 296

RESULT 16
ID Q7QPI9 GIALA PRELIMINARY; PRT; 892 AA.
AC Q7QPI9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP_173_13415_15094.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Alley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000168; EAA36927.1; -; Genomic_DNA.
DR InterPro; IPR002553; Adaptin_N.
DR Pfam; PF01602; Adaptin_N; 1.
SQ SEQUENCE 892 AA; 99003 MW; 9E78811CD1B98189 CRC64;

Query Match      75.5%; Score 37; DB 2; Length 892;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
   |:|:|:|:|
Db 593 ESARKQVMKY 602

RESULT 17
ID Q8I5C3 PLAF7 PRELIMINARY; PRT; 1132 AA.
AC Q8I5C3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PFL1430C;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22255705; PubMed=12368664; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.F., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
```

```

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.G.;
RT "Genome sequence of the human malaria parasite Plasmodium
RL falciparum.";
RL Nature 419:498-511(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
RA Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE014848; AAN36372.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1132 AA; 135030 MW; 6E8D89869A7C2FE CRC64;

Query Match 75.5%; Score 37; DB 2; Length 1132;
Best Local Similarity 70.0%; Pred. NO. 5e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQWKY 10
| ||||:|
Db 986 EKYREQVLKY 995

RESULT 18
RS13_YEAST
ID RS13_YEAST STANDARD; PRT; 150 AA.
AC P05756;
DT 01-NOV-1998 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 40S ribosomal protein S13 (S27A) (YS15).
GN Name=RPS13; Synonyms=RPS13C; OrderedLocusNames=YDR064W;
GN ORFNames=YD9609.i8, D4252;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetozoa; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=S288c;
RC MEDLINE=96381250; PubMed=8789263;
RX DOI=10.1002/(SICI)1097-0061(199601)12:1<85::AID-YEA890>3.0.CO;2-U;
RA Brandt P., Ramlow S., Otto B., Bloeker H.;
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
RN of Saccharomyces cerevisiae chromosome IV.";
RN Yeast 12:85-90(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288c / AB972;
RX MEDLINE=97313263; PubMed=9169867;
RA Barges M., Baron L., Becker A., Biteau N., Bloeker H., Blugeon C.,
RA Boskovic J., Brandt P., Bruckner M., Buhrago M.J., Coster F.,
RA Delaveau T., del Rey F., Dujon B., Eide L.G., Garcia-Cantalejo J.M.,
RA Goffeau A., Gomez-Peris A., Granotier C., Hanemann V., Haneke T.,
RA Hoheisel J.D., Jaeger W., Jimenez A., Jonniaux J.-L., Kraemer C.,
RA Kuester H., Laamanen P., Legros Y., Louis E.J., Moeller-Rieker S.,
RA Monnet A., Moro M., Mueller-Auer S., Nussbaumer B., Paricio N.,
RA Paulin L., Perez J., Perez-Alonso M., Perez-Ortin J.E., Pohl T.M.,
RA Prydz H., Purnelle B., Rasmussen S.W., Remacha M.A., Revuelta J.L.,
RA Rieger M., Salom D., Saluz H.P., Saiz J.E., Saren A.-M., Schaefer M.,
RA Scharfe M., Schmidt E.R., Schneider C., Scholler P., Schwarz S.,
RA Soler-Mira A., Urrestarazu L.A., Verhasselt P., Vissers S., Voet M.,
RA Volktaert G., Wagner G., Wambutt R., Wedler E., Wedler H., Woelfl S.,
RA Harris D.E., Bowman S., Brown D., Churcher C.M., Connor R., Dedman K.,
RA Gentles S., Hamlin N., Hunt S., Jones L., McDonald S., Murphy L.D.,
RA Niblett S., Odell C., Oliver K., Rajandream M.A., Richards C.,
RA Shore L., Walsh S.V., Barrell B.G., Dietrich F.S., Mulligan J.T.,
RA Allen E., Araujo R., Aviles E., Berno A., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Hunicke-Smith S., Hyman R.W.,
RA Komp C., Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K.,

```

```

RA Namath A., Oefner P., Oh C., Petel F.X., Roberts D., Schramm S.,
RA Schroeder M., Shogren T., Shroff N., Winant A., Yelton M.A.,
RA Botstein D., Davis R.W., Johnston M., Andrews S., Brinkman R.,
RA Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S.,
RA Greco T., Halloworth K., Hawkins J., Hillier L.W., Jier M.,
RA Johnson D., Johnston L., Kirsten J., Kucaba T., Langston Y.,
RA Latreille P., Le T., Mardis E., Menezes S., Miller N., Nhan M.,
RA Pauley A., Peluso D., Rifkin L., Riles L., Taich A., Trevasakis E.,
RA Vagnat R., Alberman K., Hani J., Heumann K., Kleine K.,
RA Wagerat R., Wohlmann P., Vaudin M., Wilson R.,
RA Mewes H.-W., Zollner A., Zaccaria P.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IV.";
RL Nature 387:75-78(1997).
RN [3]
RP PROTEIN SEQUENCE OF 1-40.
RX MEDLINE=83048950; PubMed=6814480;
RA Otake E., Higo K.-I., Osawa S.;
RT "Isolation of seventeen proteins and amino-terminal amino acid
RT sequences of eight proteins from cytoplasmic ribosomes of yeast.";
RL Biochemistry 21:4545-4550(1982).
CC -!- SIMILARITY: Belongs to the ribosomal protein S15P family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC EMBL; X84162; CAA58980.1; -; Genomic DNA.
DR EMBL; Z49209; CAA89093.1; -; Genomic DNA.
DR EMBL; Z74360; CAA98882.1; -; Genomic DNA.
DR PIR; S54048; S54048.
DR PDB; 1K5X; Model; O=65-129.
DR PDB; 1S1H; EM; O=65-129.
DR InAct; P05756; -
DR GERMOnline; 140555; -
DR Ensembl; YDR064W; Saccharomycetes cerevisiae.
DR SGD; S00002471; RPS13.
DR GO; GO:0005843; C:cytosolic small ribosomal subunit (sensu Eu. . .; TAS.
DR GO; GO:0003735; F:structural constituent of ribosome; TAS.
DR GO; GO:0006412; P:protein biosynthesis; TAS.
DR InterPro; IPR013606; Ribosomal_S13_N.
DR InterPro; IPR000589; Ribosomal_S15.
DR Pfam; PF08069; Ribosomal_S13_N; 1.
DR Pfam; PF00312; Ribosomal_S15; 1.
DR PROSITE; PS00362; RIBOSOMAL_S15; 1.
DR 3D-structure; Complete proteome; Direct protein sequencing;
KW Ribonucleoprotein; Ribosomal protein.
FT INIT MET 0
FT CONFLICT 24 24 W -> G (in Ref. 3).
FT CONFLICT 31 31 S -> C (in Ref. 3).
SQ SEQUENCE 150 AA; 16898 MW; 378D72BE81E3AF58 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 150;
Best Local Similarity 70.0%; Pred. NO. 97;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQWKY 10
| ||||:|
Db 30 ESVIEQIVKY 39

RESULT 19
Q6FT65_CANGA
ID Q6FT65_CANGA PRELIMINARY; PRT; 151 AA.
AC Q6FT65;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P05756|Saccharomycetes cerevisiae YDR064W YS15.
GN OrderedLocusNames=CAGL06050279;
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetozoa; Saccharomycetes;

```

```
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=ATCC 2901 / CBS 138;
RC PubMed=15229592; DOI=10.1038/nature02579;
RX Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sanei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L02123; AAAS3654.1; -; Genomic_DNA.
CC EMBL; U00096; AAC73885.1; -; Genomic_DNA.
CC EMBL; D90716; BAA35458.1; -; Genomic_DNA.
CC EMBL; D90717; BAA35464.1; -; Genomic_DNA.
CC PIR; F64816; F64816.
CC ECHOBASE; EB1539; -.
CC Ecogene; EG11579; ybiA.
CC InterPro; IPR012816; CHP02464 YbiA.
CC TIGRFAMs; TIGR02464; ribofla fusion; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 160 AA; 18669 MW; 28D6DA5C6748653C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 160;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMK 9
Db |||:||||:
31 ESVREQVIKY 40

RESULT 20
YBIA_ECOLI
ID YBIA_ECOLI STANDARD; PRT; 160 AA.
AC YB0176;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical protein ybiA.
GN Name=ybiA; OrderedLocNames=b0798;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=K12 / W3110;
RC PubMed=94312114; PubMed=8037924;
RX MEDLINE=94312114; PubMed=8037924;
RA Ohmori H.;
RT "Structural analysis of the rhlE gene of Escherichia coli.";
RL Jpn. J. Genet. 69:1-12(1994).
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).

[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=K12;
RC MEDLINE=97061202; PubMed=8905232;
RX Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sanei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L02123; AAAS3654.1; -; Genomic_DNA.
CC EMBL; U00096; AAC73885.1; -; Genomic_DNA.
CC EMBL; D90716; BAA35458.1; -; Genomic_DNA.
CC EMBL; D90717; BAA35464.1; -; Genomic_DNA.
CC PIR; F64816; F64816.
CC ECHOBASE; EB1539; -.
CC Ecogene; EG11579; ybiA.
CC InterPro; IPR012816; CHP02464 YbiA.
CC TIGRFAMs; TIGR02464; ribofla fusion; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 160 AA; 18669 MW; 28D6DA5C6748653C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 160;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMK 9
Db |||:||||:
90 ESVREQVMR 98

RESULT 21
Q83L08_SHIFL
ID Q83L08_SHIFL PRELIMINARY; PRT; 160 AA.
AC Q83L08; Q7C2G6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein ybiA.
GN Name=ybiA; OrderedLocNames=S0789, SF0749;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf586;
RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
```

```

RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.; comparative genomics of Shigella
RA "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; AAN42383.1; -; Genomic DNA.
DR EMBL; AE016980; AAP16260.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 160 AA; 18629 MW; AB0015667247722A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 160;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 9
|||:||||
Db 90 ESVREQVMKY 98

RESULT 22
Q50H29_BOVIN PRELIMINARY; PRT; 187 AA.
AC Q50H29;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Acetyl-CoA carboxylase, type beta (Fragment).
GN Name=acacb;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart muscle;
RA Lovstedt K., Hallen S., Osila F., Clapham J., Corneliussen B.;
RT "Isolation and characterization of cDNA encoding bovine heart acetyl-
RT CoA carboxylase type beta.";
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ966323; CAI84636.1; -; mRNA.
FT NON_TER 1
FT NON_TER 187
SQ SEQUENCE 187 AA; 21758 MW; F30C011A2705CBB5 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 187;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|||:||||
Db 52 ESVRSVMVRY 61

RESULT 23
Q4P3L0_USTMA PRELIMINARY; PRT; 193 AA.
AC Q4P3L0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
DR EMBL; U05303.1;
GN ORFNames=U05303.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelheil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A.,

```

```

RA Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer G., Gnerre S.,
RA Fitzgeraid M., Foley K., Gage D., Galagan J., Gearin H., Hafez N.,
RA Girke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Higgins H.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Iliev I.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikelsen T., Mlenga V., Moru K.,
RA Mozes J., Mullrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnaz C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tenzaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACF01000191; EAK86552.1; -; Genomic DNA.
SQ SEQUENCE 193 AA; 21243 MW; 657A60F0285CF8A6 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 193;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|||:||||
Db 85 QSIREEMLKY 94

RESULT 24
Q62DC6_BURMA PRELIMINARY; PRT; 242 AA.
AC Q62DC6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Phosphoglycolate phosphatase (EC 3.1.3.18).
GN Name=gph-2; OrderedLocNames=BMAA0534;
OS Burkholderia mallei (pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M., Dodson R.J.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouli H.M., Kolonay J.F.,

```

```

RA Madupu R., Mohammad Y., Nelson W.C., Radune D., Romero C.M.,
RA Saria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M., Burkholderia mallei genome."
RT "Structural flexibility in the Burkholderia mallei genome."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000011; AAU4683.1; -, Genomic_DNA.
DR TIGR; BMAA0534; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR006402; HAD_SF-IA-v3.
DR InterPro; IPR006439; HAD_SF_A.v1.
DR InterPro; IPR005833; Hlgase/ep_hydro.
DR InterPro; IPR006346; PGP_bact.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-v3; 1.
DR TIGRFAMs; TIGR01449; PGP_bact; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 242 AA; 25312 MW; 652EFPA29619ABD5 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 242;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 45 ETTREVMRY 54

RESULT 25
Q63QH4_BURPS PRELIMINARY; PRT; 245 AA.
ID Q63QH4_BURPS PRELIMINARY; PRT; 245 AA.
AC Q63QH4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative hydrolase.
GN OrderedLocustNames=BPSL3049;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchaveit M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH37060.1; -, Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008967; F:phosphoglycolate phosphatase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR006402; HAD_SF-IA-v3.
DR InterPro; IPR006439; HAD_SF_A.v1.
DR InterPro; IPR005833; Hlgase/ep_hydro.
DR InterPro; IPR006346; PGP_bact.
DR Pfam; PF00702; Hydrolase; 1.

```

```

DR PRINTS; PR00413; HADHALOGNASE.
DR TIGRFAMs; TIGR01549; HAD-SF-IA-v1; 1.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-v3; 1.
DR TIGRFAMs; TIGR01449; PGP_bact; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 245 AA; 25602 MW; 3FDEB0CE60E84E33 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 245;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 48 ETTREVMRY 57

RESULT 26
Q6A7W0_PROAC
ID Q6A7W0_PROAC PRELIMINARY; PRT; 418 AA.
AC Q6A7W0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tyrosyl-tRNA synthetase (EC 6.1.1.1).
GN OrderedLocustNames=PPA1404;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hoster F., Liesegang H., Wiewer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
of human skin."
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT83155.1; -, Genomic_DNA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0004831; F:tyrosine-tRNA ligase activity; IEA.
DR GO; GO:0006437; P:tyrosyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002942; S4.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; Tyr-tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTHYR.
DR TIGRFAMs; TIGR00234; TYRS; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS50889; S4; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome; Ligase.
SQ SEQUENCE 418 AA; 46257 MW; 391E0973BAD5E707 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 418;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|:|:|:|:|
Db 99 ESIKQVSKY 108

RESULT 27
Q9FH67_ARATH
ID Q9FH67_ARATH PRELIMINARY; PRT; 499 AA.
AC Q9FH67;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome P450.

```

DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN 1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane,
SQ SEQUENCE 523 AA; 59570 MW; FLD4BEF899D60E7A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 523;
Best Local Similarity 60.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Indels 0; Gaps 0

QY 1 ESRVQUMKY 10
 : |||:|:
Db 436 DEVREVLKY 445

RESULT 29
Q529X0 MAGGR
ID ID Q529X0 MAGGR PRELIMINARY; PRT; 671 AA.
AC AC Q529X0;
DT DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein.
DE ORFNames=MG02264.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
ON NCBI_TaxID=242507;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusham C., Abebe A., Abouelleil A., Adekoya E.,
RA Al-Zahrani M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhaltel B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Callimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Girke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huseby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kanat A., Kamysseis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.-J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Menus L.,
RA Mesirov J., Mihalev A., Minova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mores J., Mulrain L., Munson G., Nielsen C., Nizzari M., Norbu C.,
RA Nguyen N., Nguyen T., Nicol R., O'leary S., Omotosho B.,
RA Norbu N., O'Donnell P., Okoawo O., Perrin D., Phunkhang P., Pignani B.,
RA O'Neill K., Oseman S., Parker S., Ramasamy U., Rameau R., Raymond C.,
RA Purcell S., Rachupka T., Ramasamy U., Rodriguez J., Rogers J., Rogov P.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Sharpe T.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sougnez C.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Stavropoulos S.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S., Tchinga P.,
RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Topham K.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000439; EAA54279.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 671 AA; 77945 MW; F24B32542DGAA78 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 671;
Best Local Similarity 70.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 602 EFVRERMMKY 611

RESULT 30
Q51V60 MAGGR PRELIMINARY; PRT; 829 AA.
AC Q51V60;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG04275.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_taxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhaty B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H., Gnerre S.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kanal M., Kanat A., Kamysseilis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyttsang T., Lokyttsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Medirov J., Mihalev A., Minova T., Mikkelsen T., Mienga V., Moru K.,
RA Moses J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramaamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Statson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
```

```
RA Towe S., Teamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001087; EAA50516.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 829 AA; 96375 MW; 97C6F404C36235F8 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 829;
Best Local Similarity 70.0%; Pred. No. 5.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 740 EFVRERMMKY 749

RESULT 31
Q5CWC1 CRYPV PRELIMINARY; PRT; 970 AA.
AC Q5CWC1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Ydr449cp/Utpop; small (Ribosomal) subunit (SSU) processosome (Contains
DE U3 snRNA). HAT repeats.
GN ORFNames=cgd8 900;
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_taxID=5807;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Iowa type II;
RX PubMed=15044751; DOI=10.1126/science.1094786;
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahante J.E., Zhu G.,
RA Lanto C.A., Deng M., Liu C., Widmer G., Tripori S., Buck G.A., Xu P.,
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
RA Anantharaman V., Aravind L., Kapur V.;
RT "Complete genome sequence of the apicomplexan, Cryptosporidium
RT parvum.";
RT Science 304:441-445 (2004).
RL -1- CAUTION: The sequence shown here is derived from an
RL EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
RL preliminary data.
CC EMBL; AAE0100003; EAK89317.1; -; Genomic_DNA.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR003164; AP2_A_adaptin_C.
DR InterPro; IPR008456; Collagen_bind.
DR InterPro; IPR011992; EF-Hand_type.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001093; IMPDH/GMPRTase.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR000403; P13/_kinase_cat.
DR InterPro; IPR008925; tRNA-synt_bind.
DR InterPro; IPR009032; Vpu_cyt.
DR Pfam; PF00478; IMPDH; 1.
```



```
SQ SEQUENCE 970 AA; 113326 MW; 5D7415245601643A CRC64;
Query Match 73.5%; Score 36; DB 2; Length 970;
Best Local Similarity 60.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
||:|:|:|
Db 476 ESIREKIMLY 485

RESULT 32
Q59FY4 HUMAN PRELIMINARY; PRT; 998 AA.
ID Q59FY4
AC Q59FY4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Acetyl-Coenzyme A carboxylase alpha isoform 2 variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209325; BAD92562.1; -; mRNA.
FT NON_TER 1
SQ SEQUENCE 998 AA; 114617 MW; 78A0767D8144A97C CRC64;
Query Match 73.5%; Score 36; DB 2; Length 998;
Best Local Similarity 70.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|
Db 191 ESVRSWMVRY 200

RESULT 33
Q4SCU4 TETNG PRELIMINARY; PRT; 1182 AA.
ID Q4SCU4
AC Q4SCU4
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 7 SCAF14650, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0020333001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
```

```
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014650; CAG01538.1; -; Genomic_DNA.
FT NON_TER 1182
SQ SEQUENCE 1182 AA; 134850 MW; B7F5ECADB75F112 CRC64;
Query Match 73.5%; Score 36; DB 2; Length 1182;
Best Local Similarity 70.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|
Db 324 ESVRSWMVRY 333

RESULT 34
Q5F363 CHICK PRELIMINARY; PRT; 1233 AA.
ID Q5F363
AC Q5F363
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_32920;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chickenursal lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ851787; CAH65421.1; -; mRNA.
DR InterPro; IPR001606; ARID.
DR InterPro; IPR003347; TF_JmjC.
DR InterPro; IPR003349; TF_JmjN.
DR InterPro; IPR004198; ZnF_C5HC2.
DR Pfam; PF01388; ARID; 1.
DR Pfam; PF02373; JmjC; 1.
DR Pfam; PF02375; JmjN; 1.
DR Pfam; PF02928; zf-CSHC2; 1.
DR SMART; SM00501; BRIGHT; 1.
DR SMART; SM00558; JmjC; 1.
DR SMART; SM00545; JmjN; 1.
DR PROSITE; PS1011; ARID; 1.
KW Hypothetical protein.
SQ SEQUENCE 1233 AA; 138109 MW; 6386F8A089BBA7B7 CRC64;
Query Match 73.5%; Score 36; DB 2; Length 1233;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ESVREQVMKY 10
|||:|:|
Db 564 ESVRAQVEKY 573

RESULT 35
```

JARD2_MOUSE
ID JARD2_MOUSE STANDARD; PRT; 1234 AA.
AC Q62315; Q99LD1;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Jumonji protein (Jumonji/ARID domain-containing protein 2).
GN Name-Jarid2; Synonyms=Jmj;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC STRAIN=C57BL/6J; PubMed=7758946;
RX MEDLINE=95278734; PubMed=7758946;
RA Takeuchi T., Yamazaki Y., Katoh-Fukui Y., Tsuchiya R., Kondo S.,
RA Motoyama J., Higashinakagawa T.,
RT "Gene trap capture of a novel mouse gene, jumonji, required for neural
RT tube formation.";
RL Genes Dev. 9:1211-1222(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
RA Okazaki I., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oeato N., Saito K., Suzuki H., Yananaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schrim L.M., Kanapin A., Matzuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Meglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Negashima T., Numata K., Okido T., Pavan W.J., Pette G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Vezardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6, and Czech II; TISSUE=Brain, and Mammary gland;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bucow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huily S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=20374468; PubMed=10913339; DOI=10.1006/bbrc.2000.3138;
RA Toyoda M., Kojima M., Takeuchi T.,
RA "Jumonji is a nuclear protein that participates in the negative
RT regulation of cell growth.";
RL Biochem. Biophys. Res. Commun. 274:332-336(2000).
RN [5]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=20289961; PubMed=10807864;
RA Lee Y., Song A.J., Baker R., Micales B., Conway S.J., Lyons G.B.,
RA "Jumonji, a nuclear protein that is necessary for normal heart
RT development.";
RL Circ. Res. 86:932-938(2000).
RN [6]
RP FUNCTION, NUCLEAR LOCALIZATION SIGNAL, AND MUTAGENESIS OF
RP 106-ARG-LYS-107.
RX PubMed=12890568; DOI=10.1074/jbc.M307386200;
RA Kim T.-G., Kraus J.C., Chen J., Lee Y.,
RT "JUMONJI, a critical factor for cardiac development, functions as a
RT transcriptional repressor.";
RL J. Biol. Chem. 278:42247-42255(2003).
RN [7]
RP FUNCTION, AND INTERACTIONS WITH GATA4 AND NKX2-5.
RX PubMed=15542826; DOI=10.1128/MCB.24.23.10151-10160.2004;
RA Kim T.-G., Chen J., Sadoshima J., Lee Y.,
RT "Jumonji represses atrial natriuretic factor gene expression by
RT inhibiting transcriptional activities of cardiac transcription
RL factors.";
RL Mol. Cell. Biol. 24:10151-10160(2004).
CC -1- FUNCTION: Required for neural tube formation. Essential for normal
CC heart development and function. Acts as a transcriptional
CC repressor of ANF by binding to both GATA4 and NKX2-5 and
CC repressing their transcriptional activator activities.
CC Participates in the negative regulation of cell proliferation
CC signaling.
CC -1- SUBUNIT: Interacts with the N-terminal region of GATA4 and the C-
CC terminal region of the NKX2-5 homeobox and represses their ability
CC to activate transcription of ANF.
CC -1- INTERACTION:
CC Q8369:Gata4; NbExp=2; IntAct=EBI-493592, EBI-297008;
CC P42582:NKX2-5; NbExp=2; IntAct=EBI-493592, EBI-297021;
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Widely expressed in embryos. In adults,
CC expressed at high levels in heart, skeletal muscle, brain and
CC thymus.
CC -1- SIMILARITY: Contains 1 ARID domain.
CC -1- SIMILARITY: Contains 1 JmJC domain.
CC -1- SIMILARITY: Contains 1 JmJN domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; D31967; BAA06736.1; -; mRNA.
CC EMBL; AK045214; BAC32264.1; -; mRNA.
CC EMBL; BC003374; AAH03374.1; -; mRNA.
CC EMBL; BC052444; AAH52444.1; -; mRNA.
CC EMBL; BC060695; AAH60695.1; -; mRNA.
CC PIR; T30254; T30254.
CC IntAct; Q62315; ENSMUSG00000038518; Mus musculus.
CC Ensembl; MGI:104813; Jarid2.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003677; F:DNA binding; IDA.

```
DR GO: GO:0016564; F:transcriptional repressor activity; IDA.
DR GO: GO:0045892; P:negative regulation of transcription, DNA-d. . .; IDA.
DR InterPro: IPR001606; ARID.
DR InterPro: IPR003347; TF_JmjC.
DR InterPro: IPR003349; TF_JmjN.
DR InterPro: IPR004198; ZnF_CSHC2.
DR Pfam: PF01388; ARID; 1.
DR Pfam: PF02373; JmjC; 1.
DR Pfam: PF02375; JmjN; 1.
DR Pfam: PF02928; zf-CSHC2; 1.
DR PROSITE; PS51011; ARID; 1.
DR Developmental protein; Nuclear protein; Repressor; Transcription;
KW Transcription regulation.
FT DOMAIN 619 711 ARID
FT MOTIF 104 110 Nuclear localization signal.
FT MUTAGEN 106 107 RK->AA: Leads to cytoplasmic
FT relocalization.
FT CONFLICT 1096 1096 L -> Q (in Ref. 3; AAH03374).
SQ SEQUENCE 1234 AA; 137445 MW; B56E172C5E5745B5 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 1234;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 573 ESVRAQVEKY 582

RESULT 36
Q86LZ8 SCHMA
ID Q86LZ8 SCHMA PRELIMINARY; PRT; 1499 AA.
AC Q86LZ8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Insulin receptor protein kinase RTK-2.
GN Name=RTK-2;
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
NUCLEOTIDE SEQUENCE.
RP MEDLINE=22442601; PubMed=12554084; DOI=10.1016/S0166-6851(02)00249-9;
RX Vicogne J., Pin J.P., Lardans V., Capron M., Noel C., Dissous C.;
RT "An unusual receptor tyrosine kinase of Schistosoma mansoni contains a
RT Venus Flytrap module."
RL Mol. Biochem. Parasitol. 126:51-62(2003).
DR EMBL; AF314754; AAN39120.1; -; mRNA.
DR HSSP; P08581; 1R1W.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:005524; F:ATP binding; IEA.
DR GO: GO:0005006; P:epidermal growth factor receptor activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR000585; Hemopexin.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase AS.
DR Pfam: PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM0060; FN3; 2.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyrKG; 1.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase; Receptor.

GO: GO:0016564; F:transcriptional repressor activity; IDA.
GO: GO:0045892; P:negative regulation of transcription, DNA-d. . .; IDA.
InterPro: IPR001606; ARID.
InterPro: IPR003347; TF_JmjC.
InterPro: IPR003349; TF_JmjN.
InterPro: IPR004198; ZnF_CSHC2.
Pfam: PF01388; ARID; 1.
Pfam: PF02373; JmjC; 1.
Pfam: PF02375; JmjN; 1.
Pfam: PF02928; zf-CSHC2; 1.
PROSITE; PS51011; ARID; 1.
Developmental protein; Nuclear protein; Repressor; Transcription;
Transcription regulation.
DOMAIN 619 711 ARID
MOTIF 104 110 Nuclear localization signal.
MUTAGEN 106 107 RK->AA: Leads to cytoplasmic
relocalization.
CONFLICT 1096 1096 L -> Q (in Ref. 3; AAH03374).
SEQUENCE 1234 AA; 137445 MW; B56E172C5E5745B5 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 1234;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 573 ESVRAQVEKY 582

RESULT 36
Q86LZ8 SCHMA
ID Q86LZ8 SCHMA PRELIMINARY; PRT; 1499 AA.
AC Q86LZ8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Insulin receptor protein kinase RTK-2.
GN Name=RTK-2;
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
NUCLEOTIDE SEQUENCE.
RP MEDLINE=22442601; PubMed=12554084; DOI=10.1016/S0166-6851(02)00249-9;
RX Vicogne J., Pin J.P., Lardans V., Capron M., Noel C., Dissous C.;
RT "An unusual receptor tyrosine kinase of Schistosoma mansoni contains a
RT Venus Flytrap module."
RL Mol. Biochem. Parasitol. 126:51-62(2003).
DR EMBL; AF314754; AAN39120.1; -; mRNA.
DR HSSP; P08581; 1R1W.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:005524; F:ATP binding; IEA.
DR GO: GO:0005006; P:epidermal growth factor receptor activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR000494; EGFR_L.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR006212; Furin repeat.
DR InterPro: IPR000585; Hemopexin.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase AS.
DR Pfam: PF01030; Recep_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM0060; FN3; 2.
DR SMART; SM00261; FU; 1.
DR SMART; SM00219; TyrKG; 1.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase; Receptor.

SQ SEQUENCE 1499 AA; 170354 MW; A603560611E26582 CRC64;
Query Match 73.5%; Score 36; DB 2; Length 1499;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SVREQVMKY 10
| | | | |
Db 176 SVREQIVRY 184

RESULT 37
Q8IDM0 PLAF7
ID Q8IDM0 PLAF7 PRELIMINARY; PRT; 1847 AA.
AC Q8IDM0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein MAL13P1.239.
GN Name=MAL13P1.239;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Ackin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the ENBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52601.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1847 AA; 221588 MW; DA999B6FE8DF25A CRC64;

Query Match 73.5%; Score 36; DB 2; Length 1847;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
| | | | |
Db 854 ENTREQIMVY 863

RESULT 38
COAC_CHICK
ID COAC_CHICK STANDARD; PRT; 2324 AA.
AC P11029;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) [includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Name=ACAC;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=88139305; PubMed=2893793;
RA Takai T., Yokoyama C., Wada K., Tanabe T.;
RT "Primary structure of chicken liver acetyl-CoA carboxylase deduced
RT from cDNA sequence."
RL J. Biol. Chem. 263:2651-2657(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 493-820.
RC TISSUE=Liver;
RX MEDLINE=87106011; PubMed=2879745; DOI=10.1016/0014-5793(87)81564-8;
RA Takai T., Wada K., Tanabe T.;
RT "Primary structure of the biotin-binding site of chicken liver acetyl-
RT CoA carboxylase.";
```

RL FEBS Lett. 212:98-102(1987).
CC -|- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC -|- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
CC + malonyl-CoA.
CC -|- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -|- COFACTOR: Biotin.
CC -|- ENZYME REGULATION: By phosphorylation.
CC -|- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Contains 1 ATP-grasp domain.
CC -|- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -|- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -|- SIMILARITY: Contains 1 carboxyltransferase domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; J03541; AAA48701.1; -; mRNA.
CC EMBL; X05019; CAA28675.1; -; mRNA.
CC PIR; A29924; A29924.
CC HSSP; Q00955; 10D4.
CC Ensembl; ENSGALG00000005439; Gallus gallus.
CC InterPro; IPR011761; ATP_GRASP.
CC InterPro; IPR011764; BC.
CC InterPro; IPR001882; Biotin_BS.
CC InterPro; IPR005482; Biotin_carb_C.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR000022; Carboxyl_trans.
CC InterPro; IPR011763; COA_CT_C.
CC InterPro; IPR011762; COA_CT_N.
CC InterPro; IPR005481; Cpsase_L_N.
CC InterPro; IPR005479; Cpsase_D2_ATP_bd.
CC Pfam; PF02785; Biotin_carb_C; 1.
CC Pfam; PF00364; Biotin_lipoyl; 1.
CC Pfam; PF01039; Carboxyl_trans; 1.
CC Pfam; PF00289; Cpsase_L_chain; 1.
CC Pfam; PF02786; Cpsase_L_D2; 1.
CC PROSITE; PS09975; ATP_GRASP; 1.
CC PROSITE; PS09979; BC; 1.
CC PROSITE; PS00188; BIOTIN; 1.
CC PROSITE; PS09968; BIOTINYL_LIPOYL; 1.
CC PROSITE; PS09989; COA_CT_CTER; 1.
CC PROSITE; PS09980; COA_CT_NTER; 1.
CC PROSITE; PS00866; CPSASE 1; 1.
CC PROSITE; PS00867; CPSASE 2; 1.
CC ATP-binding; Biotin; Direct protein sequencing;
KW Fatty acid biosynthesis; Ligase; Lipid synthesis;
KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1675 2171 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).
FT ACT_SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent).
FT BINDING 1800 1800 Coenzyme A (By similarity).
FT BINDING 2104 2104 Coenzyme A (By similarity).
FT BINDING 2106 2106 Coenzyme A (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1193 1193 Phosphoserine (By similarity).
SQ SEQUENCE 2324 AA; 262720 MW; 3F1C541F01BBBFF6 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2324;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ESUREQVMKY 10
| | | | | | | |
Db 1467 ESVRSMVMRY 1476

RESULT 39
COAL_RAT STANDARD; PRT; 2345 AA.
AC P11497; P97902;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin
DE carboxylase (EC 6.3.4.14)].
GN Name=Acaca; Synonyms=Acac;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=88320328; PubMed=2901088;
RA Lopez-Casillas F., Bai D.-H., Luo X., Kong I.-S., Hermodson M.A.,
RA Kim K.-H.;
RT "Structure of the coding sequence and primary amino acid sequence of
RT acetyl-coenzyme A carboxylase.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5784-5788(1988).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=89264558; PubMed=2566999;
RA Luo X.N., Park K., Lopez-Casillas F., Kim K.-H.;
RT "Structural features of the acetyl-CoA carboxylase gene: mechanisms
RT for the generation of mRNAs with 5' end heterogeneity.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4042-4046(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-33.
RX MEDLINE=89214151; PubMed=2565337;
RA Lopez-Casillas F., Kim K.-H.;
RT "Heterogeneity at the 5' end of rat acetyl-coenzyme A carboxylase
RT mRNA. Lipogenic conditions enhance synthesis of a unique mRNA in
RT liver.";
RL J. Biol. Chem. 264:7176-7184(1989).
RN [4]
RP NUCLEOTIDE SEQUENCE OF 1167-1200 (ISOFORMS 1 AND 2), AND
RP PHOSPHORYLATION SITE SER-1200.
RX MEDLINE=90337981; PubMed=1974251;
RA Kong I.-S., Lopez-Casillas F., Kim K.-H.;
RT "Acetyl-CoA carboxylase mRNA species with or without inhibitory coding
RT sequence for Ser-1200 phosphorylation.";
RL J. Biol. Chem. 265:13695-13701(1990).
RN [5]
RP BIOTIN-BINDING SITE
RX MEDLINE=89289706; PubMed=2567668;
RA Bai D.-H., Moon T.-W., Lopez-Casillas F., Andrews P.C., Kim K.-H.;
RT "Analysis of the biotin-binding site on acetyl-CoA carboxylase from
RT rat.";
RL Eur. J. Biochem. 182:239-245(1989).
RN [6]
RP PROTEIN SEQUENCE OF 76-85 AND 1198-1201, AND PHOSPHORYLATION.
RX MEDLINE=88296498; PubMed=2900138;
RA Munday M.R., Campbell D.G., Carling D., Hardie D.G.;
RT "Identification by amino acid sequencing of three major regulatory
RT phosphorylation sites on rat acetyl-CoA carboxylase.";
RL Eur. J. Biochem. 175:331-338(1988).
RN [7]
RP PARTIAL PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RX STRAIN=Wiistar; TISSUE=Liver;
RX MEDLINE=94237850; PubMed=7910165;
RA Winz R., Hess D., Aebersold R., Brownsey R.W.;
RT "Unique structural features and differential phosphorylation of the

RT 280-kDa component (isozyme) of rat liver acetyl-CoA carboxylase.";
RL J. Biol. Chem. 269:14438-14445(1994).
CC -!- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis
CC of long-chain fatty acids. This protein carries three functions:
CC biotin carboxyl carrier protein, biotin carboxylase, and
CC carboxyltransferase.
CC -!- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
CC + malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- COFACTOR: Biotin.
CC -!- ENZYME REGULATION: By phosphorylation.
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting)
CC step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P11497-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P11497-2; Sequence=VSP_011753;
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Contains 1 ATP-grasp domain.
CC -!- SIMILARITY: Contains 1 biotin carboxylation domain.
CC -!- SIMILARITY: Contains 1 biotinyl-binding domain.
CC -!- SIMILARITY: Contains 1 carboxyltransferase domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; J03808; AAA40653.1; -; mRNA.
CC EMBL; M26731; AAA40652.1; -; Genomic_DNA.
CC EMBL; M26195; AAA40654.1; -; mRNA.
CC EMBL; M26196; AAA40655.1; -; mRNA.
CC EMBL; M26197; AAA40656.1; -; mRNA.
CC EMBL; M55315; -; NOT_ANNOTATED_CDS; mRNA.
CC PIR; A35578; A35578.
CC HSP; Q00955; I0D4.
CC Ensembl; ENSRNOG0000034013; Rattus norvegicus.
CC RGD; 621248; Acac.
CC InterPro; IPR011761; ATP GRASP.
CC InterPro; IPR011764; BC.
CC InterPro; IPR001882; Biotin BS.
CC InterPro; IPR005482; Biotin carb C.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR000022; Carboxyl_trans.
CC InterPro; IPR011763; COA CT C.
CC InterPro; IPR011762; COA CT N.
CC InterPro; IPR005479; CPase D2 ATP_bd.
CC Pfam; PF02785; Biotin_carb_C; 1.
CC Pfam; PF00364; Biotin_lipoyl; 1.
CC Pfam; PF01039; Carboxyl_trans; 1.
CC Pfam; PF00289; CPase L_chain; 1.
CC Pfam; PF02786; CPase L_D2; 1.
CC PROSITE; PS00975; ATP GRASP; 1.
CC PROSITE; PS00979; BC; 1.
CC PROSITE; PS00188; BIOTIN; 1.
CC PROSITE; PS00968; BIOTINYL_LIPOYL; 1.
CC PROSITE; PS00989; COA CT CTER; 1.
CC PROSITE; PS00980; COA CT NTER; 1.
CC PROSITE; PS00866; CPASE_1; 1.
CC PROSITE; PS00867; CPASE_2; 1.
CC
CC Alternative splicing; ATP-binding; Biotin; Direct protein sequencing;
CC Fatty acid biosynthesis; Ligase; Lipid synthesis;
CC Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
CC
CC DOMAIN 116 617 Biotin carboxylation.
CC DOMAIN 274 465 ATP-grasp.
CC DOMAIN 751 817 Biotinyl/lipoyl.
CC DOMAIN 1697 2193 Carboxyltransferase.

FT NP_BIND 314 319 ATP (Potential).
FT ACT_SITE 440 440 By similarity.
FT BINDING 785 785 Biotin (covalent).
FT BINDING 1822 1822 Coenzyme A (By similarity).
FT BINDING 2126 2126 Coenzyme A (By similarity).
FT BINDING 2128 2128 Coenzyme A (By similarity).
FT MOD_RES 77 77 Phosphoserine.
FT MOD_RES 79 79 Phosphoserine.
FT MOD_RES 1200 1200 Phosphoserine.
FT VARSPIC 1189 1196 Missing (in isoform 2).
FT FTID=VSP_011753 /FTID=VSP_011753
SQ SEQUENCE 2345 AA; 265194 MW; 78E9CF9ADE1E8771 CRC64;
Query Match 73.5%; Score 36; DB 1; Length 2345;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
Db 1489 ESVRSMVMRY 1498
RESULT 40
Q5SWU9 MOUSE
ID Q5SWU9 MOUSE PRELIMINARY; PRT; 2345 AA.
AC Q5SWU9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Acetyl-Coenzyme A carboxylase.
GN Name=Acac; ORFNames=RP23-123010.1-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Matthews N.
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Tracey A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596447; CAI25271.1; -; Genomic_DNA.
DR EMBL; AL596252; CAI24019.1; -; Genomic_DNA.
DR EMBL; AL596252; CAI25271.1; JOINED; Genomic_DNA.
DR EMBL; AL596447; CAI24019.1; JOINED; Genomic_DNA.
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR011761; ATP GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin BS.
DR InterPro; IPR005482; Biotin carb C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA CT C.
DR InterPro; IPR011762; COA CT N.
DR InterPro; IPR005479; Cphap synth_L_D2.
DR Pfam; PF02785; Biotin carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPase L_chain; 1.
DR Pfam; PF02786; CPase L_D2; 1.
DR PROSITE; PS00975; ATP GRASP; 1.
DR PROSITE; PS00979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00968; BIOTINYL_LIPOYL; 1.

KW Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
FT DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 275 466 ATP-grasp.
FT DOMAIN 752 818 Biotinyl/lipoyl.
FT DOMAIN 1698 2194 Carboxyltransferase.
FT NP_BIND 315 320 ATP (Potential).
FT ACT_SITE 441 441 By similarity.
FT BINDING 786 786 Biotin (covalent) (By similarity).
FT BINDING 1823 2127 Coenzyme A (By similarity).
FT BINDING 2127 2127 Coenzyme A (By similarity).
FT BINDING 2129 2129 Coenzyme A (By similarity).
FT MOD_RES 29 29 Phosphoserine (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1201 1201 Phosphoserine (By similarity).
SQ SEQUENCE 2346 AA; 265303 MW; 32886C5D03EEAE0E CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|||||
Db 1490 ESVRSVMWRY 1499

RESULT 43
COAL HUMAN STANDARD; PRT; 2346 AA.
ID Q013085;
AC Q13085;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 13-SEP-2005 (Rel. 48; Last annotation update)
DE Acetyl-CoA carboxylase 1 (EC 6.4.1.2) (ACC-alpha) [Includes: Biotin carboxylase (EC 6.3.4.14)].
GN Name=ACACA; Synonyms=ACAC, ACC1, ACCA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95249602; PubMed=7732023;
RA Abu-Elheiga L., Jayakumar A., Baldini A., Chirala S.S., Wakil S.J.;
RT "Human acetyl-CoA carboxylase: characterization, molecular cloning, and evidence for two isoforms";
RL Proc. Natl. Acad. Sci. U.S.A. 92:4011-4015(1995).
RN [2]
RP PHOSPHORYLATION SITE SER-29.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J., Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC !- FUNCTION: Catalyzes the rate-limiting reaction in the biogenesis of long-chain fatty acids. This protein carries three functions: biotin carboxyl carrier protein, biotin carboxylase, and carboxyltransferase.
CC !- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3) (-) = ADP + phosphate + malonyl-CoA.
CC !- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2) = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC !- COFACTOR: Biotin.
CC !- ENZYME REGULATION: By phosphorylation (By similarity).
CC !- PATHWAY: Long-chain fatty acid biosynthesis; first (rate-limiting) step.
CC !- SUBCELLULAR LOCATION: Cytoplasmic.
CC !- TISSUE SPECIFICITY: Expressed in brain, placental, skeletal muscle, renal, pancreatic and adipose tissues; expressed at low level in pulmonary tissue; not detected in the liver.
CC !- SIMILARITY: Contains 1 ATP-grasp domain.

CC !- SIMILARITY: Contains 1 biotin carboxylation domain.
CC !- SIMILARITY: Contains 1 biotinyl-binding domain.
CC !- SIMILARITY: Contains 1 carboxyltransferase domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
CC EMBL; U19822; AAC50139.1; -; mRNA.
DR PIR; I38928; I38928.
DR HSP; Q00955; I0D4.
DR Ensembl; ENSG00000132142; Homo sapiens.
DR HGNC; HGNC:84; ACACA.
DR MIM; 200350; -.
DR GO; GO:0003989; F:acetyl-CoA carboxylase activity; TAS.
DR InterPro; IPR011763; ATP_GRASP.
DR InterPro; IPR011764; BC.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR011763; COA_CT_C.
DR InterPro; IPR011762; COA_CT_N.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR005479; CPase_D2_ATP_bd.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF0289; CPase_L_D2; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR PROSITE; PS0975; ATP_GRASP; 1.
DR PROSITE; PS0979; BC; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS0968; BIOTIN_LIPOYL; 1.
DR PROSITE; PS0989; COA_CT_CTER; 1.
DR PROSITE; PS0980; COA_CT_NTER; 1.
DR PROSITE; PS00866; CPASE 1; 1.
DR PROSITE; PS00867; CPASE 2; 1.
KW ATP-binding; Biotin; Fatty acid biosynthesis; Ligase; Lipid synthesis; Multifunctional enzyme; Nucleotide-binding; Phosphorylation.
KW DOMAIN 117 618 Biotin carboxylation.
FT DOMAIN 117 618 ATP-grasp.
FT DOMAIN 275 466 Biotinyl/lipoyl.
FT DOMAIN 752 818 Carboxyltransferase.
FT DOMAIN 1698 2194 ATP (Potential).
FT NP_BIND 315 320 By similarity.
FT ACT_SITE 441 441 Biotin (covalent) (By similarity).
FT BINDING 786 786 Coenzyme A (By similarity).
FT BINDING 1823 2127 Coenzyme A (By similarity).
FT BINDING 2127 2127 Coenzyme A (By similarity).
FT BINDING 2129 2129 Phosphoserine (By similarity).
FT MOD_RES 29 29 Phosphoserine (By similarity).
FT MOD_RES 78 78 Phosphoserine (By similarity).
FT MOD_RES 80 80 Phosphoserine (By similarity).
FT MOD_RES 1201 1201 Phosphoserine (By similarity).
SQ SEQUENCE 2346 AA; 265040 MW; 9519190D40190D14 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2346;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ESVREQVMKY 10
|||||
Db 1490 ESVRSVMWRY 1499

RESULT 44
COAL SHEEP STANDARD; PRT; 2346 AA.
ID Q28559;
AC Q28559;
DT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)

AC	Q9BS18;
AD	01-JUN-2001 (TrEMBLrel. 17, Created)
AE	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
AF	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
AG	Anaphase promoting complex subunit 13.
AN	Name=ANAPC13;
AO	Homo sapiens (Human).
AP	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AQ	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
AR	Homo.
AS	NCBI_TaxID=9606;
AT	NCBI_TaxID=9606;
AV	11;
AW	NUCLEOTIDE SEQUENCE.
AX	TISSUE=Kidney;
AY	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
AZ	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
BA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
BB	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
BC	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
BD	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
BE	Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
BF	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,
BG	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
BH	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
BI	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
BJ	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
BK	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
BL	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
BM	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
BN	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
BO	Butterfield Y.S.N., Krzywiecki M.I., Skalska U., Smalley D.E.,
BP	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
BQ	"generation and initial analysis of more than 15,000 full-length human
BR	RT and mouse cDNA sequences.";
BS	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
BT	[2]
BU	NUCLEOTIDE SEQUENCE.
BV	TISSUE=Kidney;
BW	NTM MGC Project;
BX	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
BY	EMBL; BC005398; AAH05398.1; -; mRNA.
BZ	Ensembl; ENSG00000129055; Homo sapiens.
CA	SEQUENCE 74 AA; 8521 MW; F3C52A7E3821CB3B CRC64;
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY</	

RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CR858529; CAH90756.1; -, mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 74 AA; 8507 MW; F3C52A7E39AD18B CRC64;
 Query Match 71.4%; Score 35; DB 2; Length 74;
 Best Local Similarity 70.0%; Pred. No. 72;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ESVREQWVKY 10
 |||||
 Db 48 ESVREQEMKW 57
 |||||
 RESULT 49
 Q8R034_MOUSE
 ID Q8R034_MOUSE PRELIMINARY; PRT; 74 AA.
 AC Q8R034;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Anapc13 protein (Mus musculus 10 day old male pancreas cDNA, RIKEN
 full-length enriched library, clone:1810004D07 product:HYPOTHETICAL
 8.5 kbp PROTEIN homolog).
 DE Name=Anapc13;
 GN Name=Anapc13;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSU=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schut N.K.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.G.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Haeh P.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.B.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSU=Mammary gland;
 RX Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSU=Pancreas;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSU=Pancreas;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

SAITO T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seyta T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSU=Pancreas;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,
 Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
 Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 Konagaya A., Kurochkin I.V., Lee Y., Lenhard R.M., Lyons P.A.,
 Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 Verardo R., Wagner L., Wehstedt C., Wang Y., Watanabe Y., Wells C.,
 Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
 Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSU=Pancreas;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSU=Pancreas;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto T., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).

[8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Pancreas;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028526; AAH28526.1; -; mRNA.
DR EMBL; AK007332; BAC25171.1; -; mRNA.
DR Ensemble; ENSMUSG000000305048; Mus musculus.
DR MGI; MGI:1916260; Anapc13.
KW Hypothetical protein.
SQ SEQUENCE 74 AA; 9346 MW; E84F16A13821CB3B CRC64;
Query Match 71.4%; Score 35; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 72;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ESVREQVMKY 10
|||:||||:
Db 48 ESKVEQEMKW 57
Search completed: May 12, 2006, 10:50:56
Job time : 106.769 secs

RESULT 50
Q4SKC6 TETNG
ID Q4SKC6_TETNG PRELIMINARY; PRT; 74 AA.
AC Q4SKC6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 13 SCAP14566, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00016792001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.-P., Guzy J.,
RA Farra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014566; CAF98906.1; -; Genomic_DNA.
FT NON_TER 74

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 87.1795 Seconds
(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTLEWLHGSPC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq 21:*

1: Genesep1980a:*
2: Genesep1990a:*
3: Genesep2000a:*
4: Genesep2001a:*
5: Genesep2002a:*
6: Genesep2003a:*
7: Genesep2003bs:*
8: Genesep2004a:*
9: Genesep2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	100.0	12	AAW09524	Aaw09524 Thrombopo
2	78	100.0	12	AAW36675	Aaw36675 Thrombopo
3	78	100.0	12	AAW16993	Aaw16993 TPO-mimet
4	78	100.0	12	AAW25894	Aaw25894 Human thr
5	78	100.0	12	ABB72879	Abb72879 TPO mimet
6	78	100.0	12	ADJ73030	Adj73030 TPO mimet
7	78	100.0	12	ADJ52665	Adj52665 CH1 delet
8	78	100.0	12	AAW09464	Aaw09464 Thrombopo
9	78	100.0	16	AAW33329	Aaw33329 Thrombopo
10	78	100.0	16	AAW17019	Aaw17019 TPO-mimet
11	78	100.0	16	AAW25829	Aaw25829 Human thr
12	78	100.0	16	ABB72905	Abb72905 TPO mimet
13	78	100.0	16	ADJ73057	Adj73057 TPO mimet
14	78	100.0	16	ADJ52692	Adj52692 CH1 delet
15	78	100.0	16	ADJ51653	Adj51653 CH1 delet
16	78	100.0	16	AAW09582	Aaw09582 Thrombopo
17	59	75.6	12	AAW36733	Aaw36733 Thrombopo
18	59	75.6	12	AAW25952	Aaw25952 Human thr
19	59	75.6	12	AAW09572	Aaw09572 Thrombopo
20	54	69.2	12	AAW36723	Aaw36723 Thrombopo
21	54	69.2	12	AAW16994	Aaw16994 TPO-mimet
22	54	69.2	12	AAW25942	Aaw25942 Human thr
23	54	69.2	12	ABB72880	Abb72880 TPO mimet
24	54	69.2	12	ABB72880	Abb72880 TPO mimet

25	54	69.2	12	7	ADJ73031	Adj73031 TPO mimet
26	54	69.2	12	8	ADJ52666	Adj52666 CH1 delet
27	54	69.2	12	8	ADJ51627	Adj51627 CH1 delet
28	52	66.7	12	2	AAW09581	Aaw09581 Thrombopo
29	52	66.7	12	4	AAU25951	Aau25951 Human thr
30	52	66.7	13	2	AAW36732	Aaw36732 Thrombopo
31	49	62.8	10	2	AAW09566	Aaw09566 Thrombopo
32	49	62.8	10	2	AAW36717	Aaw36717 Thrombopo
33	49	62.8	10	4	AAU25936	Aau25936 Human thr
34	48	61.5	12	3	AAW17002	Aaw17002 TPO-mimet
35	48	61.5	12	5	ABB72888	Abb72888 TPO mimet
36	48	61.5	12	8	ADJ52674	Adj52674 CH1 delet
37	48	61.5	12	8	ADJ51635	Adj51635 CH1 delet
38	47	60.3	12	2	AAW09570	Aaw09570 Thrombopo
39	47	60.3	12	2	AAW36721	Aaw36721 Thrombopo
40	47	60.3	12	4	AAU25940	Aau25940 Human thr
41	46	59.0	10	2	AAW09507	Aaw09507 Thrombopo
42	46	59.0	10	2	AAW36658	Aaw36658 Thrombopo
43	46	59.0	10	3	AAW16990	Aaw16990 TPO-mimet
44	46	59.0	10	4	AAU25877	Aau25877 Human thr
45	46	59.0	10	5	ABB72876	Abb72876 TPO mimet
46	46	59.0	10	7	ADJ73027	Adj73027 TPO mimet
47	46	59.0	10	8	ADJ52662	Adj52662 CH1 delet
48	46	59.0	10	8	ADJ51623	Adj51623 CH1 delet
49	46	59.0	12	2	AAW09523	Aaw09523 Thrombopo
50	46	59.0	12	2	AAW36674	Aaw36674 Thrombopo
51	46	59.0	12	4	AAU25893	Aau25893 Human thr
52	46	59.0	12	4	AAW87994	Aaw87994 Rice abio
53	45	57.7	10	2	AAW09512	Aaw09512 Thrombopo
54	45	57.7	10	2	AAW36663	Aaw36663 Thrombopo
55	45	57.7	10	3	AAW16992	Aaw16992 TPO-mimet
56	45	57.7	10	4	AAU25882	Aau25882 Human thr
57	45	57.7	10	5	ABB72878	Abb72878 TPO mimet
58	45	57.7	10	5	ABB73029	Abb73029 TPO mimet
59	45	57.7	10	8	ADJ52664	Adj52664 CH1 delet
60	45	57.7	10	8	ADJ51625	Adj51625 CH1 delet
61	45	57.7	12	2	AAW09571	Aaw09571 Thrombopo
62	45	57.7	12	2	AAW36763	Aaw36763 Thrombopo
63	45	57.7	12	2	AAW36722	Aaw36722 Thrombopo
64	45	57.7	12	4	AAU25941	Aau25941 Human thr
65	45	57.7	12	4	AAW09494	Aaw09494 Thrombopo
66	45	57.7	19	2	AAW09461	Aaw09461 Thrombopo
67	45	57.7	19	2	AAW33028	Aaw33028 Thrombopo
68	45	57.7	19	2	AAW36645	Aaw36645 Thrombopo
69	45	57.7	19	3	AAW17022	Aaw17022 TPO-mimet
70	45	57.7	19	4	AAU25864	Aau25864 Human thr
71	45	57.7	19	4	AAU25825	Aau25825 Human thr
72	45	57.7	19	5	ABB72908	Abb72908 TPO mimet
73	45	57.7	19	7	ADJ73060	Adj73060 TPO mimet
74	45	57.7	19	8	ADJ52695	Adj52695 CH1 delet
75	45	57.7	19	8	ADJ51656	Adj51656 CH1 delet
76	45	57.7	53	2	AAW02934	Aaw02934 Fragment
77	45	57.7	53	7	ADA07767	Ada07767 Human sec
78	45	57.7	53	8	ADN41280	Adn41280 Novel hum
79	45	57.7	532	8	ADN19273	Adn19273 Bacterial
80	44.5	57.1	18	2	AAW09589	Aaw09589 Thrombopo
81	44.5	57.1	18	2	AAW36740	Aaw36740 Thrombopo
82	44.5	57.1	18	4	AAU25959	Aau25959 Human thr
83	44	56.4	12	2	AAW09577	Aaw09577 Thrombopo
84	44	56.4	12	2	AAW36728	Aaw36728 Thrombopo
85	44	56.4	12	4	AAU25947	Aau25947 Human thr
86	44	56.4	13	2	AAW09603	Aaw09603 Thrombopo
87	44	56.4	13	2	AAW36754	Aaw36754 Thrombopo
88	44	56.4	13	4	AAU25973	Aau25973 Human thr
89	44	56.4	18	7	ADN59672	Adn59672 Thrombopo
90	44	56.4	22	7	ADN59839	Adn59839 Tmp pept1
91	44	56.4	25	7	ADN59744	Adn59744 Thrombopo
92	43	55.1	9	7	ADJ73040	Adj73040 TPO mimet
93	43	55.1	9	7	ADJ73037	Adj73037 TPO mimet
94	43	55.1	9	7	ADJ73041	Adj73041 TPO mimet
95	43	55.1	9	7	ADJ73039	Adj73039 TPO mimet
96	43	55.1	9	7	ADJ73038	Adj73038 TPO mimet
97	43	55.1	10	3	AAW17000	Aaw17000 TPO-mimet

98	43	55.1	10	5	ABB72886	Abb72886 TPO mimet	171	41	52.6	619	4	AAU33715	Aau33715 Pseudomon
99	43	55.1	10	8	ADJ52672	ADJ52672 CH1 delet	172	41	52.6	619	6	ABUI5512	Abui5512 Protein e
100	43	55.1	10	8	ADJ51633	ADJ51633 CH1 delet	173	41	52.6	770	3	AAy54060	Aay54060 Amino aci
101	43	55.1	11	3	ABAB17001	Abab17001 TPO-mimet	174	41	52.6	776	3	AAy54062	Aay54062 Amino aci
102	43	55.1	11	5	ABB72887	Abb72887 TPO mimet	175	41	52.6	804	7	ABO75043	AbO75043 Pseudomon
103	43	55.1	11	8	ADJ52673	ADJ52673 CH1 delet	176	41	52.6	826	4	ABG21979	Abg21979 Novel hum
104	43	55.1	11	8	ADJ51634	ADJ51634 CH1 delet	177	41	52.6	833	3	AAy54059	Aay54059 Amino aci
105	43	55.1	13	3	ABAB17003	Abab17003 TPO-mimet	178	41	52.6	839	3	AAy54061	Aay54061 Amino aci
106	43	55.1	13	5	ABB72889	Abb72889 TPO mimet	179	41	52.6	853	3	AAy54058	Aay54058 Amino aci
107	43	55.1	13	8	ADJ52675	ADJ52675 CH1 delet	180	41	52.6	853	8	ADU06633	AdU06633 Novel bro
108	43	55.1	13	8	ADJ51636	ADJ51636 CH1 delet	181	41	52.6	859	3	AAy54056	Aay54056 Amino aci
109	43	55.1	14	3	ABAB17004	Abab17004 TPO-mimet	182	41	52.6	1025	5	ABR90967	AbR90967 Herbicida
110	43	55.1	14	5	ABB72890	Abb72890 TPO mimet	183	41	52.6	1025	8	ADRO4093	AdR04093 A thalian
111	43	55.1	14	8	ADJ52676	ADJ52676 CH1 delet	184	41	52.6	1043	6	ABUI5501	Abui5501 Protein e
112	43	55.1	14	8	ADJ51637	ADJ51637 CH1 delet	185	41	52.6	1151	7	ABO74850	AbO74850 Pseudomon
113	43	55.1	18	2	AAW09583	Aaw09583 Thrombopo	186	41	52.6	4342	4	AAU33611	Aau33611 Pseudomon
114	43	55.1	18	2	AAW36734	Aaw36734 Thrombopo	187	41	52.6	4342	6	ABUI5500	Abui5500 Protein e
115	43	55.1	18	4	AAU25953	Aau25953 Human thr	188	40.5	51.9	3588	8	ADQ39441	Adq39441 Human myo
116	43	55.1	583	5	ABP65849	Abp65849 Bifidobac	189	40.5	51.9	4346	8	ADQ39440	Adq39440 Human myo
117	43	55.1	791	6	AAE36091	Aae36091 Human dia	190	40.5	51.9	4347	8	ADQ39439	Adq39439 Human myo
118	43	55.1	791	6	AAE36092	Aae36092 Human dia	191	40.5	51.9	4370	7	ADJ69461	Adj69461 Human hea
119	43	55.1	791	7	ABE33701	AbE33701 Human 348	192	40.5	51.9	4391	6	AAE34390	Aae34390 Human per
120	42	53.8	12	2	AAW09568	Aaw09568 Thrombopo	193	40.5	51.9	4393	4	AAAB31889	Aab31889 Amino aci
121	42	53.8	12	2	AAW09569	Aaw09569 Thrombopo	194	40.5	51.9	4393	8	ADL35758	AdL35758 Human per
122	42	53.8	12	2	AAW36719	Aaw36719 Thrombopo	195	40.5	51.9	4393	8	ADQ39442	Adq39442 Human myo
123	42	53.8	12	2	AAW36720	Aaw36720 Thrombopo	196	40.5	51.9	4436	4	ABG23265	Abg23265 Novel hum
124	42	53.8	12	4	AAU25939	Aau25939 Human thr	197	40	51.3	12	2	AAW09525	Aaw09525 Thrombopo
125	42	53.8	12	4	AAU25938	Aau25938 Human thr	198	40	51.3	12	2	AAW09529	Aaw09529 Thrombopo
126	42	53.8	14	2	AAW09541	Aaw09541 Thrombopo	199	40	51.3	12	2	AAW09528	Aaw09528 Thrombopo
127	42	53.8	14	2	AAW36770	Aaw36770 Thrombopo	200	40	51.3	12	2	AAW09567	Aaw09567 Thrombopo
128	42	53.8	14	3	AAW36692	Aaw36692 Thrombopo	201	40	51.3	12	2	AAW09526	Aaw09526 Thrombopo
129	42	53.8	14	3	ABAB1699	Abab1699 TPO-mimet	202	40	51.3	12	2	AAW36676	Aaw36676 Thrombopo
130	42	53.8	14	4	AAU25911	Aau25911 Human thr	203	40	51.3	12	2	AAW36677	Aaw36677 Thrombopo
131	42	53.8	14	5	ABB72885	Abb72885 TPO mimet	204	40	51.3	12	2	AAW36718	Aaw36718 Thrombopo
132	42	53.8	14	7	ADJ73036	ADJ73036 TPO mimet	205	40	51.3	12	2	AAW36718	Aaw36718 Thrombopo
133	42	53.8	14	8	ADJ52671	ADJ52671 CH1 delet	206	40	51.3	12	2	AAW36680	Aaw36680 Thrombopo
134	42	53.8	14	8	ADJ51632	ADJ51632 CH1 delet	207	40	51.3	12	4	AAU25895	Aau25895 Human thr
135	42	53.8	18	2	AAW09596	Aaw09596 Thrombopo	208	40	51.3	12	4	AAU25937	Aau25937 Human thr
136	42	53.8	18	2	AAW36747	Aaw36747 Thrombopo	209	40	51.3	12	4	AAU25896	Aau25896 Human thr
137	42	53.8	18	2	AAU25966	Aau25966 Human thr	210	40	51.3	12	4	AAU25898	Aau25898 Human thr
138	42	53.8	18	4	AAU25967	Aau25967 Human thr	211	40	51.3	12	4	AAU25899	Aau25899 Human thr
139	42	53.8	18	4	ABO59773	AbO59773 Human gen	212	40	51.3	13	3	AAAB16995	Aab16995 TPO-mimet
140	42	53.8	44	8	ABO59773	AbO59773 Human gen	213	40	51.3	13	5	ABB72881	Abb72881 TPO mimet
141	42	53.8	56	4	ABG00428	Abg00428 Novel hum	214	40	51.3	13	7	ADJ73032	Adj73032 TPO mimet
142	42	53.8	173	9	ADY28796	Ady28796 CD47 bind	215	40	51.3	13	8	ADJ52667	Adj52667 CH1 delet
143	42	53.8	224	9	ABO71269	AbO71269 Pseudomon	216	40	51.3	13	8	ADJ51628	Adj51628 CH1 delet
144	42	53.8	287	9	ABE33606	Aeb33606 L. pneumo	217	40	51.3	14	2	AAW09536	Aaw09536 Thrombopo
145	42	53.8	287	9	ABE33618	Aeb33618 L. pneumo	218	40	51.3	14	2	AAW36687	Aaw36687 Thrombopo
146	42	53.8	1200	4	ABB63533	Abb63533 Drosophil	219	40	51.3	14	2	AAW36768	Aaw36768 Thrombopo
147	42	53.8	1481	4	ABG13697	Abg13697 Novel hum	220	40	51.3	14	4	AAU25906	Aau25906 Human thr
148	41.5	53.2	257	8	ADU21782	AdU21782 Bacterial	221	40	51.3	18	2	AAW09592	Aaw09592 Thrombopo
149	41	52.6	10	2	AAW09547	Aaw09547 Thrombopo	222	40	51.3	18	2	AAW09595	Aaw09595 Thrombopo
150	41	52.6	10	2	AAW36698	Aaw36698 Thrombopo	223	40	51.3	18	2	AAW36746	Aaw36746 Thrombopo
151	41	52.6	10	4	AAU25917	Aau25917 Human thr	224	40	51.3	18	2	AAW36743	Aaw36743 Thrombopo
152	41	52.6	12	2	AAW09580	Aaw09580 Thrombopo	225	40	51.3	18	4	AAU25965	Aau25965 Human thr
153	41	52.6	12	2	AAW36731	Aaw36731 Thrombopo	226	40	51.3	18	4	AAU25962	Aau25962 Human thr
154	41	52.6	12	4	AAU25950	Aau25950 Human thr	227	40	51.3	66	3	RAG53141	Rag53141 Arabidops
155	41	52.6	218	8	ADR99369	AdR99369 Chlamydia	228	40	51.3	67	7	ADC27609	AdC27609 Human col
156	41	52.6	242	2	AAy35100	Aay35100 Chlamydia	229	40	51.3	73	4	ABG00429	Abg00429 Novel hum
157	41	52.6	306	7	ADM26482	Adm26482 Hyperther	230	40	51.3	77	8	ADX68706	Adx68706 Plant ful
158	41	52.6	310	8	ADY55980	Ady55980 Plant pol	231	40	51.3	80	4	AAU17763	Aau17763 Novel hum
159	41	52.6	486	6	ABP77210	Abp77210 N. gonorr	232	40	51.3	80	7	ADG41143	Adg41143 Human res
160	41	52.6	489	4	AAAB92701	AaB92701 Human pro	233	40	51.3	80	7	ADI96917	Adi96917 Human res
161	41	52.6	513	2	AAy38789	Aay38789 Neisseria	234	40	51.3	140	8	ADP80867	Adp80867 Mouse par
162	41	52.6	513	2	AAy38786	Aay38786 Neisseria	235	40	51.3	169	7	ABM86829	Abm86829 Rice abio
163	41	52.6	513	2	AAy38787	Aay38787 Neisseria	236	40	51.3	176	5	AAE25747	Aae25747 Rice Ft h
164	41	52.6	513	2	AAy38788	Aay38788 Neisseria	237	40	51.3	176	8	ADR04263	AdR04263 Rice Ft h
165	41	52.6	513	9	ABE49392	Aeb49392 N. mening	238	40	51.3	231	6	ABU94266	Abu94266 Lolium pe
166	41	52.6	513	9	ABE49394	Aeb49394 N. mening	239	40	51.3	286	3	RAG68809	Rag68809 Arabidops
167	41	52.6	513	9	ABE49395	Aeb49395 N. gonorr	240	40	51.3	364	8	ADX89766	Adx89766 Plant ful
168	41	52.6	513	9	ABE49397	Aeb49397 N. gonorr	241	40	51.3	380	7	ADE56536	Ade56536 Rat Prote
169	41	52.6	604	6	ADA48262	Ada48262 Rice prot	242	40	51.3	408	6	ABU33917	Abu33917 Protein e
170	41	52.6	604	7	ABM86602	Abm86602 Rice abio	243	40	51.3	480	3	AAG06808	Aag06808 Arabidops

244	40	51.3	483	9	ABM93529	Abm93529 M. xanthu	317	39	50.0	1472	7	ADE57585	Ades7585 Rat Prote
245	40	51.3	510	3	AG06807	Aag06807 Arabidops	318	39	50.0	1472	7	ADD47124	Aded47124 Rat Prote
246	40	51.3	775	8	AD93666	Adg93666 Plant ful	319	39	50.0	1472	7	ADE63422	Aded63422 Rat Prote
247	40	51.3	1567	4	ABB64992	Abb64992 Drosophil	320	39	50.0	1472	7	AD83343	Aded83343 Rat Prote
248	40	51.3	1872	2	AAW68510	Aaw68510 Partial h	321	39	50.0	1472	7	AEb22116	Aeb22116 Rat alpha
249	40	51.3	2000	8	ADP23496	Adp23496 PRO polyp	322	39	50.0	1474	2	AAR11334	Aar11334 Recombina
250	40	51.3	2180	8	ADN22356	Adn22356 Bacterial	323	39	50.0	1474	3	AAy97157	Aay97157 Human alp
251	40	51.3	2318	8	ADJ76296	Adj76296 Marker ge	324	39	50.0	1474	3	AAy97157	Aay97157 Human alp
252	40	51.3	2321	2	AAW48698	Aaw48698 Human Not	325	39	50.0	1474	5	AAU81017	Aau81017 Human alp
253	40	51.3	2321	2	AAW48698	Aaw48698 Human Not	326	39	50.0	1474	5	AAU74798	Aau74798 Human alp
254	40	51.3	2321	8	ADH34621	Adh34621 Notch hom	327	39	50.0	1474	5	ABP65218	Abp65218 Hypoxia-r
255	40	51.3	2321	8	ADJ75570	Adj75570 Marker ge	328	39	50.0	1474	7	AAO31219	Aao31219 Human alp
256	40	51.3	2321	8	ADN48701	Adn48701 Human Not	329	39	50.0	1474	7	AAO31224	Aao31224 Human alp
257	40	51.3	2321	8	ADN48701	Adn48701 Human Not	330	39	50.0	1474	7	AAO31221	Aao31221 Human alp
258	39	50.0	2321	9	ADX70405	Adx70405 Human Not	331	39	50.0	1474	7	AAO31222	Aao31222 Human alp
259	39	50.0	16	2	AAW09609	Aaw09609 Thrombopo	332	39	50.0	1474	7	ADe63424	Aded63424 Human Pro
260	39	50.0	16	2	AAW36760	Aaw36760 Thrombopo	333	39	50.0	1474	7	ADe63424	Aded63424 Human Pro
261	39	50.0	16	4	AAU25979	Aau25979 Human thr	334	39	50.0	1474	7	ADe57586	Aded57586 Human Pro
262	39	50.0	51	8	ADU52159	Adu52159 Human sec	335	39	50.0	1474	7	ADd47130	Aded47130 Human Pro
263	39	50.0	58	8	ADS52157	Ads52157 Human sec	336	39	50.0	1474	7	ADd47126	Aded47126 Human Pro
264	39	50.0	67	4	AAW16701	Aam16701 Peptide #	337	39	50.0	1474	7	ADd47130	Aded47130 Human Pro
265	39	50.0	67	4	ABB35684	Abb35684 Peptide #	338	39	50.0	1474	7	ADd47130	Aded47130 Human Pro
266	39	50.0	67	4	AAW29188	Aam29188 Peptide #	339	39	50.0	1474	7	ADd47130	Aded47130 Human Pro
267	39	50.0	67	4	ABB21115	Abb21115 Protein #	340	39	50.0	1474	8	ADQ39717	Adq39717 Human myo
268	39	50.0	67	4	AAW68879	Aam68879 Human bon	341	39	50.0	1474	8	ADQ39717	Adq39717 Human myo
269	39	50.0	67	4	AAW68879	Aam68879 Human bon	342	39	50.0	1474	9	ADW96863	Adw96863 Human nas
270	39	50.0	67	4	ABG50540	Abg50540 Human liv	343	39	50.0	1474	9	ABW09519	Abw09519 Human pre
271	39	50.0	67	5	ABG38458	Abg38458 Human pep	344	39	50.0	1479	8	ADQ10075	Ado10075 Novel hum
272	39	50.0	113	5	ABW08067	Abw08067 Human alp	345	39	50.0	1482	7	ADd47134	Aded47134 Human Pro
273	39	50.0	174	9	ADY28813	Ady28813 CD4 bindi	346	39	50.0	1482	7	ADf90915	Adf90915 Human hep
274	39	50.0	183	5	ABW08068	Abw08068 Human alp	347	39	50.0	1482	7	ADd47134	Aded47134 Human Pro
275	39	50.0	186	2	AAW08067	Aay08067 Rat serin	348	39	50.0	1500	4	AAO31218	Aao31218 Human alp
276	39	50.0	190	8	ADN24475	Adn24475 Bacterial	349	39	50.0	1542	4	ABG28005	Abg28005 Novel hum
277	39	50.0	192	5	ABP05718	Abp05718 Human ORF	350	39	50.0	1542	7	ABG28005	Abg28005 Novel hum
278	39	50.0	201	8	ADN21721	Adn21721 Bacterial	351	39	50.0	1543	7	ABG28005	Abg28005 Novel hum
279	39	50.0	206	6	ABU20005	Abu20005 Protein e	352	39	50.0	2531	5	ADi16935	Adi16935 Rat NOVX
280	39	50.0	206	6	ABU21268	Abu21268 Protein e	353	39	50.0	2531	5	ADi16935	Adi16935 Rat NOVX
281	39	50.0	218	7	ABO78209	Abu78209 Pseudomon	354	39	50.0	2531	5	ADi16934	Adi16934 Rat NOVX
282	39	50.0	254	6	AAU39399	Abu39399 Protein e	355	39	50.0	2531	7	ADe63705	Aded63705 Rat Prote
283	39	50.0	260	2	AAW41989	Aaw41989 Flea seri	356	39	50.0	2531	7	ADe63705	Aded63705 Rat Prote
284	39	50.0	260	4	ABW50622	Aab50622 Flea seri	357	39	50.0	2531	7	ADe63709	Aded63709 Rat Prote
285	39	50.0	288	2	AAW34527	Aaw34527 hTCP prot	358	39	50.0	2531	7	ADe63701	Aded63701 Rat Prote
286	39	50.0	288	2	AAW73931	Aaw73931 T-lymphoc	359	39	50.0	2531	8	ADo58587	Ado58587 Mouse Not
287	39	50.0	288	2	AAW18077	Aay18077 CD4+ T-ly	360	39	50.0	2531	8	ADZ12495	Adz12495 Murine ca
288	39	50.0	307	3	AAW31478	Aag31478 Arabidops	361	39	50.0	4317	6	ABU39938	Abu39938 Protein e
289	39	50.0	323	9	ADY866742	Ady866742 Sbfl rest	362	39	50.0	4317	6	ABU39938	Abu39938 Protein e
290	39	50.0	323	9	ADY866742	Ady866742 Sbfl rest	363	39	50.0	4317	6	ABU39938	Abu39938 Protein e
291	39	50.0	339	9	ADY53330	Ady53330 Streptomy	364	39	50.0	4317	6	ABU39938	Abu39938 Protein e
292	39	50.0	358	7	ADW09375	Adw09375 Novel pro	365	39	50.0	4317	6	ABU39938	Abu39938 Protein e
293	39	50.0	375	7	ADW09375	Adw09375 Novel pro	366	39	50.0	4317	6	ABU39938	Abu39938 Protein e
294	39	50.0	384	6	ABW53688	Abw53688 Protein s	367	39	50.0	4317	6	ABU39938	Abu39938 Protein e
295	39	50.0	384	7	ADK64056	Adk64056 Disease t	368	39	50.0	4317	6	ABU39938	Abu39938 Protein e
296	39	50.0	450	8	ADX92629	Adx92629 plant ful	369	39	50.0	4317	6	ABU39938	Abu39938 Protein e
297	39	50.0	492	5	ABW31476	Aag31476 Arabidops	370	39	50.0	4317	6	ABU39938	Abu39938 Protein e
298	39	50.0	492	5	ABW2817	Abb2817 Herbicida	371	39	50.0	4317	6	ABU39938	Abu39938 Protein e
299	39	50.0	799	2	AAW73934	Aaw73934 T-lymphoc	372	39	50.0	4317	6	ABU39938	Abu39938 Protein e
300	39	50.0	799	2	AAW18073	Aay18073 CD4+ T-ly	373	39	50.0	4317	6	ABU39938	Abu39938 Protein e
301	39	50.0	841	8	ABW82746	Abw82746 Human dia	374	39	50.0	4317	6	ABU39938	Abu39938 Protein e
302	39	50.0	842	6	ABU11716	Abu11716 Human MDD	375	39	50.0	4317	6	ABU39938	Abu39938 Protein e
303	39	50.0	881	2	AAW73933	Aaw73933 T-lymphoc	376	39	50.0	4317	6	ABU39938	Abu39938 Protein e
304	39	50.0	881	2	AAW18072	Aay18072 CD4+ T-ly	377	39	50.0	4317	6	ABU39938	Abu39938 Protein e
305	39	50.0	913	7	ADP65188	Adp65188 Human pro	378	39	50.0	4317	6	ABU39938	Abu39938 Protein e
306	39	50.0	915	8	ADU64366	Adu64366 Human pro	379	39	50.0	4317	6	ABU39938	Abu39938 Protein e
307	39	50.0	915	8	AAW52306	Aaw52306 Human pro	380	39	50.0	4317	6	ABU39938	Abu39938 Protein e
308	39	50.0	915	2	AAW73932	Aaw73932 T-lymphoc	381	39	50.0	4317	6	ABU39938	Abu39938 Protein e
309	39	50.0	915	2	AAW18071	Aay18071 CD4+ T-ly	382	39	50.0	4317	6	ABU39938	Abu39938 Protein e
310	39	50.0	1285	3	ABW43949	Abw43949 Human can	383	39	50.0	4317	6	ABU39938	Abu39938 Protein e
311	39	50.0	1450	7	ADK41536	Adk41536 Anti-cell	384	39	50.0	4317	6	ABU39938	Abu39938 Protein e
312	39	50.0	1450	9	ADW96864	Adw96864 Human mat	385	39	50.0	4317	6	ABU39938	Abu39938 Protein e
313	39	50.0	1451	5	AAU81018	Aau81018 Human alp	386	39	50.0	4317	6	ABU39938	Abu39938 Protein e
314	39	50.0	1451	7	ADK41537	Adk41537 Anti-cell	387	39	50.0	4317	6	ABU39938	Abu39938 Protein e
315	39	50.0	1472	7	ADd47128	Aded47128 Rat Prote	388	39	50.0	4317	6	ABU39938	Abu39938 Protein e
316	39	50.0	1472	7	ADe60379	Aded60379 Rat Prote	389	39	50.0	4317	6	ABU39938	Abu39938 Protein e

390	38	48.7	94	8	ADK47669	Adk47669 Streptoco	463	38	48.7	500	6	ABR98776	Human sec
391	38	48.7	104	3	RAY87158	Ray87158 Human sec	464	38	48.7	500	6	ABO16299	Human sec
392	38	48.7	104	4	AAE06135	Aae06135 Human gen	465	38	48.7	500	6	ABR92199	Human sec
393	38	48.7	104	5	ABG33958	Abg33958 Human sec	466	38	48.7	500	6	ABO18840	Human sec
394	38	48.7	241	9	ABE40561	Ab40561 L. pneumo	467	38	48.7	500	6	ABR78261	Human sec
395	38	48.7	246	8	ADX94643	Adx94643 Plant ful	468	38	48.7	500	6	ABU84997	Novel hum
396	38	48.7	250	9	ABE37237	Ab37237 L. pneumo	469	38	48.7	500	6	ABO00136	Novel hum
397	38	48.7	257	5	ABG69047	Abg69047 Amino aci	470	38	48.7	500	6	ABO11468	Human sec
398	38	48.7	257	8	ADN27256	Adn27256 Bacterial	471	38	48.7	500	6	ABO02113	Human sec
399	38	48.7	284	4	ABG15780	Abg15780 Novel hum	472	38	48.7	500	6	ABU88687	Novel hum
400	38	48.7	300	4	AAAM39907	Aam39907 Human pol	473	38	48.7	500	6	ABU83382	Human sec
401	38	48.7	300	4	AAU08687	Aau08687 Human FCT	474	38	48.7	500	6	ABO06183	Novel hum
402	38	48.7	300	7	ADB32041	Adb32041 Human FCT	475	38	48.7	500	6	ABR59219	Human sec
403	38	48.7	338	7	ADE55432	Ade55432 Rat Prote	476	38	48.7	500	6	ABO09281	Human sec
404	38	48.7	338	8	ADQ20762	Adq20762 Human sof	477	38	48.7	500	6	ABO19145	Novel hum
405	38	48.7	339	2	AAW74896	Aaw74896 Human sec	478	38	48.7	500	6	ABO11163	Human sec
406	38	48.7	339	5	ABG95348	Abg95348 Human nov	479	38	48.7	500	6	ABR66781	Human sec
407	38	48.7	339	6	ABO34542	Ab34542 Region of	480	38	48.7	500	6	ABO15994	Human sec
408	38	48.7	339	7	ADI23203	Adi23203 Novel hum	481	38	48.7	500	6	ABO13700	Human sec
409	38	48.7	339	8	ADH74205	Adh74205 Human sec	482	38	48.7	500	6	ABU84926	Human sec
410	38	48.7	355	8	ADS22232	Ads22232 Bacterial	483	38	48.7	500	6	ABU55603	Human sec
411	38	48.7	360	3	AAAG31508	Aag31508 Arabidops	484	38	48.7	500	6	ABO07451	Human PRO
412	38	48.7	370	6	ADAS5399	Ada5399 Human pro	485	38	48.7	500	6	ABO03638	Human sec
413	38	48.7	370	6	ABR41179	Ab41179 Human DIT	486	38	48.7	500	6	ABR67086	Human sec
414	38	48.7	371	5	ABP69844	Abp69844 Human pol	487	38	48.7	500	6	ABO15689	Human sec
415	38	48.7	372	8	ADP24070	Adp24070 PRO polyp	488	38	48.7	500	6	ABU55970	Human sec
416	38	48.7	379	3	ASG47664	Asg47664 Arabidops	489	38	48.7	500	6	ABU61124	Human PRO
417	38	48.7	389	4	ABG13806	Abg13806 Novel hum	490	38	48.7	500	6	ABU65298	Human PRO
418	38	48.7	394	3	AAO31507	Aao31507 Arabidops	491	38	48.7	500	6	ABU95243	Novel hum
419	38	48.7	412	7	ABO69766	Ab69766 Pseudomon	492	38	48.7	500	6	ABU71146	Human PRO
420	38	48.7	413	3	ABG47663	Abg47663 Arabidops	493	38	48.7	500	6	ABO07756	Human PRO
421	38	48.7	421	8	ADS27311	Ads27311 Bacterial	494	38	48.7	500	6	ABR69997	Human sec
422	38	48.7	431	3	AB44816	Ab44816 Human ORF	495	38	48.7	500	6	ABR69330	Human sec
423	38	48.7	432	8	ADS26952	Ads26952 Bacterial	496	38	48.7	500	6	ABO01471	Human PRO
424	38	48.7	432	8	ADS26578	Ads26578 Bacterial	497	38	48.7	500	6	ABU81273	Human PRO
425	38	48.7	445	5	ABBN1957	Abbn1957 Herbicida	498	38	48.7	500	6	ABR60070	Human sec
426	38	48.7	461	3	ABG29255	Abg29255 Arabidops	499	38	48.7	500	6	ABR67805	Human sec
427	38	48.7	482	3	ABG31506	Abg31506 Arabidops	500	38	48.7	500	6	ABR65193	Human sec
428	38	48.7	482	5	ABBN90925	Abbn90925 Herbicida	501	38	48.7	500	6	ABR68415	Human sec
429	38	48.7	494	4	AAU58735	Aau58735 Propionib	502	38	48.7	500	6	ABR71827	Human sec
430	38	48.7	494	6	ABM55254	Abm55254 Propionib	503	38	48.7	500	6	ABU85307	Human PRO
431	38	48.7	500	2	ABY41738	Ab41738 Human PRO	504	38	48.7	500	6	ABU88997	Human PRO
432	38	48.7	500	3	AB44294	Ab44294 Human PRO	505	38	48.7	500	6	ABU83077	Human sec
433	38	48.7	500	3	AB447999	Ab447999 Cocoa cri	506	38	48.7	500	6	ABU94933	Novel hum
434	38	48.7	500	4	AAU29058	Aau29058 Human PRO	507	38	48.7	500	6	ABU90481	Novel hum
435	38	48.7	500	5	ABG90983	Abg90983 Breast sp	508	38	48.7	500	6	ABU83992	Human sec
436	38	48.7	500	6	ABU58434	Abu58434 Human PRO	509	38	48.7	500	6	ABU93643	Novel hum
437	38	48.7	500	6	ABU87982	Abu87982 Novel hum	510	38	48.7	500	6	ABR64888	Human sec
438	38	48.7	500	6	ABU84297	Abu84297 Human sec	511	38	48.7	500	6	ABR68720	Human sec
439	38	48.7	500	6	ABR66171	Ab66171 Human sec	512	38	48.7	500	6	ABO06536	Human sec
440	38	48.7	500	6	ABR65561	Ab65561 Human sec	513	38	48.7	500	6	ABR99081	Human sec
441	38	48.7	500	6	ABU99501	Abu99501 Human sec	514	38	48.7	500	6	ABU56965	Human PRO
442	38	48.7	500	6	ABU82740	Abu82740 Human PRO	515	38	48.7	500	6	ABU85917	Novel hum
443	38	48.7	500	6	ABU89861	Abu89861 Novel hum	516	38	48.7	500	6	ABU832204	Novel hum
444	38	48.7	500	6	ABR68110	Ab68110 Human sec	517	38	48.7	500	6	ABU87215	Human PRO
445	38	48.7	500	6	ABR56795	Ab56795 Human sec	518	38	48.7	500	6	ABU83687	Human sec
446	38	48.7	500	6	ABU96163	Abu96163 Novel hum	519	38	48.7	500	6	ABO08061	Human PRO
447	38	48.7	500	6	ABU92594	Abu92594 Human sec	520	38	48.7	500	6	ABU81772	Novel hum
448	38	48.7	500	6	ABO08671	Ab008671 Human sec	521	38	48.7	500	6	ABU65936	Novel hum
449	38	48.7	500	6	ABO02723	Ab002723 Human sec	522	38	48.7	500	6	ABR59765	Human sec
450	38	48.7	500	6	ABR74877	Ab74877 Human sec	523	38	48.7	500	6	ABU93953	Novel hum
451	38	48.7	500	6	ABR94639	Ab94639 Human sec	524	38	48.7	500	6	ABU80393	Human sec
452	38	48.7	500	6	ABO25240	Ab025240 Novel hum	525	38	48.7	500	6	ABU99806	Novel hum
453	38	48.7	500	6	ABU85612	Abu85612 Human PRO	526	38	48.7	500	6	ABR66476	Human sec
454	38	48.7	500	6	ABU98772	Abu98772 Novel hum	527	38	48.7	500	6	ABR90894	Human sec
455	38	48.7	500	6	ABU97987	Abu97987 Novel hum	528	38	48.7	500	6	ABU94321	Human PRO
456	38	48.7	500	6	ABU91693	Abu91693 Novel hum	529	38	48.7	500	6	ABU79203	Human PRO
457	38	48.7	500	6	ABU72246	Abu72246 Novel hum	530	38	48.7	500	6	ABU86532	Human sec
458	38	48.7	500	6	ABU89386	Abu89386 Human PRO	531	38	48.7	500	6	ABU86837	Novel hum
459	38	48.7	500	6	ABU86227	Abu86227 Human sec	532	38	48.7	500	6	ABU94626	Human PRO
460	38	48.7	500	6	ABU67440	Abu67440 Human sec	533	38	48.7	500	6	ABO04553	Human PRO
461	38	48.7	500	6	ABU80468	Abu80468 Human PRO	534	38	48.7	500	6	ABR70302	Human sec
462	38	48.7	500	6	ABR99386	Ab99386 Human sec	535	38	48.7	500	6	ABU98467	Human PRO

536	38	48.7	500	6	ABR65866	ABr-65866 Human sec	609	38	48.7	500	6	ABO02418	ABO02418 Human sec
537	38	48.7	500	6	ABR64583	ABr-64583 Human sec	610	38	48.7	500	6	ABR90589	ABr90589 Human sec
538	38	48.7	500	6	ABU79508	ABu79508 Human PRO	611	38	48.7	500	6	ABR73657	ABr73657 Human sec
539	38	48.7	500	6	ABU92899	ABu92899 Human sec	612	38	48.7	500	6	ABO16909	ABO16909 Human sec
540	38	48.7	500	6	ABU95858	ABu95858 Human PRO	613	38	48.7	500	6	ABR94334	ABr94334 Human sec
541	38	48.7	500	6	ABU91078	ABu91078 Novel hum	614	38	48.7	500	6	ABR75841	ABr75841 Human sec
542	38	48.7	500	6	ABU90171	ABu90171 Novel hum	615	38	48.7	500	6	ABR71217	ABr71217 Human sec
543	38	48.7	500	6	ABO09586	ABO09586 Human sec	616	38	48.7	500	6	ABR93114	ABr93114 Human sec
544	38	48.7	500	6	ABO10858	ABO10858 Human sec	617	38	48.7	500	6	ABR93419	ABr93419 Human sec
545	38	48.7	500	6	ABR70912	ABr70912 Human PRO	618	38	48.7	500	6	ABR87844	ABr87844 Human sec
546	38	48.7	500	6	ABU87520	ABu87520 Human PRO	619	38	48.7	500	6	ABO29979	ABO29979 Human sec
547	38	48.7	500	6	ABU91388	ABu91388 Human PRO	620	38	48.7	500	6	ABO33188	ABO33188 Human PRO
548	38	48.7	500	6	ABU84602	ABu84602 Human sec	621	38	48.7	500	6	ABM04876	ABM04876 Human sec
549	38	48.7	500	6	ABR69692	ABr69692 Human sec	622	38	48.7	500	6	ABM08836	ABM08836 Human sec
550	38	48.7	500	6	ABU80069	ABu80069 Human PRO	623	38	48.7	500	6	ABO36436	ABO36436 Human sec
551	38	48.7	500	6	ABU93338	ABu93338 Human PRO	624	38	48.7	500	6	ABO35521	ABO35521 Human PRO
552	38	48.7	500	6	ABO09891	ABO09891 Human sec	625	38	48.7	500	6	ABO39486	ABO39486 Human sec
553	38	48.7	500	6	ABO08976	ABO08976 Human sec	626	38	48.7	500	6	ABO35521	ABO35521 Human PRO
554	38	48.7	500	6	ABU10544	ABu10544 Human sec	627	38	48.7	500	6	ABM10361	ABM10361 Human sec
555	38	48.7	500	6	ABU95553	ABu95553 Human PRO	628	38	48.7	500	6	ABM11886	ABM11886 Human sec
556	38	48.7	500	6	ABU96762	ABu96762 Novel hum	629	38	48.7	500	6	ABO52032	ABO52032 Human PRO
557	38	48.7	500	6	ABR70607	ABr70607 Human sec	630	38	48.7	500	6	ABO52337	ABO52337 Human PRO
558	38	48.7	500	6	ABO04958	ABO04958 Novel hum	631	38	48.7	500	6	ABO23655	ABO23655 Human sec
559	38	48.7	500	6	ABO08366	ABO08366 Human sec	632	38	48.7	500	6	ABR97141	ABr97141 Human sec
560	38	48.7	500	6	ABO05573	ABO05573 Human sec	633	38	48.7	500	6	ABR86929	ABr86929 Human sec
561	38	48.7	500	6	ABR73362	ABr73362 Human sec	634	38	48.7	500	6	ABM10971	ABM10971 Human sec
562	38	48.7	500	6	ABR95554	ABr95554 Human sec	635	38	48.7	500	6	ABM28115	ABM28115 Human sec
563	38	48.7	500	6	ABR80851	ABr80851 Human sec	636	38	48.7	500	6	ABO32114	ABO32114 Human sec
564	38	48.7	500	6	ABR81156	ABr81156 Human sec	637	38	48.7	500	6	ABM15241	ABM15241 Human sec
565	38	48.7	500	6	ABM00852	ABM00852 Human sec	638	38	48.7	500	6	ABM06396	ABM06396 Human sec
566	38	48.7	500	6	ABR88454	ABr88454 Human sec	639	38	48.7	500	6	ABM04207	ABM04207 Human sec
567	38	48.7	500	6	ABM77275	ABm77275 Human sec	640	38	48.7	500	6	ABM22320	ABM22320 Human sec
568	38	48.7	500	6	ABO28759	ABO28759 Human sec	641	38	48.7	500	6	ABM07616	ABM07616 Human sec
569	38	48.7	500	6	ABO31504	ABO31504 Human sec	642	38	48.7	500	6	ABO40706	ABO40706 Human sec
570	38	48.7	500	6	ABM07921	ABm07921 Human sec	643	38	48.7	500	6	ABM33116	ABM33116 Human sec
571	38	48.7	500	6	ABO40401	ABO40401 Human sec	644	38	48.7	500	6	ABO52642	ABO52642 Human PRO
572	38	48.7	500	6	ABO35826	ABO35826 Human PRO	645	38	48.7	500	6	ABO50202	ABO50202 Human sec
573	38	48.7	500	6	ABO43965	ABO43965 Human sec	646	38	48.7	500	6	ABU99196	ABu99196 Human sec
574	38	48.7	500	6	ADA77822	Ada77822 Human sec	647	38	48.7	500	6	ABO04248	ABO04248 Human sec
575	38	48.7	500	6	ABR24760	ABr24760 Human sec	648	38	48.7	500	6	ABO05878	ABO05878 Human sec
576	38	48.7	500	6	ABO03028	ABO03028 Human sec	649	38	48.7	500	6	ABM18418	ABM18418 Human sec
577	38	48.7	500	6	ABR90284	ABr90284 Human sec	650	38	48.7	500	6	ABR97446	ABr97446 Human sec
578	38	48.7	500	6	ABM17198	ABm17198 Human sec	651	38	48.7	500	6	ABR80546	ABr80546 Human sec
579	38	48.7	500	6	ABR94944	ABr94944 Human sec	652	38	48.7	500	6	ABM01157	ABM01157 Human sec
580	38	48.7	500	6	ABR95249	ABr95249 Human sec	653	38	48.7	500	6	ABR88759	ABr88759 Human sec
581	38	48.7	500	6	ABR97751	ABr97751 Human sec	654	38	48.7	500	6	ABM13411	ABM13411 Human sec
582	38	48.7	500	6	ABR7539	ABr7539 Human sec	655	38	48.7	500	6	ABM20795	ABM20795 Human sec
583	38	48.7	500	6	ABR7580	ABr7580 Human sec	656	38	48.7	500	6	ABO41926	ABO41926 Human sec
584	38	48.7	500	6	ABM27810	ABm27810 Human sec	657	38	48.7	500	6	ABO42536	ABO42536 Human sec
585	38	48.7	500	6	ABM06091	ABm06091 Human sec	658	38	48.7	500	6	ABM10056	ABM10056 Human sec
586	38	48.7	500	6	ABM03597	ABm03597 Human sec	659	38	48.7	500	6	ABO38571	ABO38571 Human sec
587	38	48.7	500	6	ABM35048	ABm35048 Human sec	660	38	48.7	500	6	ABM32811	ABM32811 Human sec
588	38	48.7	500	6	ABM26285	ABm26285 Human sec	661	38	48.7	500	6	ABM22625	ABM22625 Human sec
589	38	48.7	500	6	ABM26285	ABm26285 Human sec	662	38	48.7	500	6	ABM74836	ABM74836 Human sec
590	38	48.7	500	6	ABO48067	ABO48067 Human sec	663	38	48.7	500	6	ADA79614	Ada79614 Human sec
591	38	48.7	500	6	ABR2809	ABr2809 Human sec	664	38	48.7	500	6	ADA24902	Ada24902 Novel hum
592	38	48.7	500	6	ABO24570	ABO24570 Human sec	665	38	48.7	500	6	ABR96226	ABr96226 Human sec
593	38	48.7	500	6	ABM11581	ABm11581 Human sec	666	38	48.7	500	6	ABR86624	ABr86624 Human sec
594	38	48.7	500	6	ABM02682	ABM02682 Human sec	667	38	48.7	500	6	ABR62317	ABr62317 Human sec
595	38	48.7	500	6	ABM15978	ABm15978 Human sec	668	38	48.7	500	6	ABR86319	ABr86319 Human sec
596	38	48.7	500	6	ABO27539	ABO27539 Human sec	669	38	48.7	500	6	ABM16588	ABM16588 Human sec
597	38	48.7	500	6	ABM29030	ABm29030 Human sec	670	38	48.7	500	6	ABM29640	ABM29640 Human sec
598	38	48.7	500	6	ABM07006	ABM07006 Human sec	671	38	48.7	500	6	ABO29064	ABO29064 Human sec
599	38	48.7	500	6	ABM21100	ABm21100 Human sec	672	38	48.7	500	6	ABM23845	ABM23845 Human sec
600	38	48.7	500	6	ABM09446	ABM09446 Human sec	673	38	48.7	500	6	ABM23235	ABM23235 Human sec
601	38	48.7	500	6	ABO41316	ABO41316 Human sec	674	38	48.7	500	6	ABO37656	ABO37656 Human sec
602	38	48.7	500	6	ABO36131	ABO36131 Human PRO	675	38	48.7	500	6	ABM22015	ABM22015 Human sec
603	38	48.7	500	6	ABO43660	ABO43660 Human PRO	676	38	48.7	500	6	ABO37656	ABO37656 Human sec
604	38	48.7	500	6	ABM76360	ABm76360 Human sec	677	38	48.7	500	6	ABM28420	ABM28420 Human sec
605	38	48.7	500	6	ABM76056	ABm76056 Human sec	678	38	48.7	500	6	ABM28725	ABM28725 Human sec
606	38	48.7	500	6	ABM25675	ABM25675 Human sec	679	38	48.7	500	6	ABM66369	ABM66369 Human sec
607	38	48.7	500	6	ABM25980	ABM25980 Human sec	680	38	48.7	500	6	ABM75751	ABM75751 Human sec
608	38	48.7	500	6	ABO03333	ABO03333 Human sec	681	38	48.7	500	6	ABM34031	ABM34031 Human sec

682	38	48.7	500	6	ABM34336	Human	sec	Abm34336	Human	sec	755	38	48.7	500	6	ABR94029	Human	sec	AbR94029	Human	sec
683	38	48.7	500	6	ABO19695	Novel	hum	AbO19695	Novel	hum	756	38	48.7	500	6	ABR79936	Human	sec	AbR79936	Human	sec
684	38	48.7	500	6	ABO20267	Human	sec	AbO20267	Human	sec	757	38	48.7	500	6	ABM11276	Human	sec	AbM11276	Human	sec
685	38	48.7	500	6	ABO21182	Human	sec	AbO21182	Human	sec	758	38	48.7	500	6	ABO32883	Human	sec	AbO32883	Human	sec
686	38	48.7	500	6	ABO22097	Human	sec	AbO22097	Human	sec	759	38	48.7	500	6	ABO30589	Human	sec	AbO30589	Human	sec
687	38	48.7	500	6	ADA12563	Human	sec	AdA12563	Human	sec	760	38	48.7	500	6	ABO30894	Human	sec	AbO30894	Human	sec
688	38	48.7	500	6	ABR96531	Human	sec	AbR96531	Human	sec	761	38	48.7	500	6	ABM27200	Human	sec	AbM27200	Human	sec
689	38	48.7	500	6	ABR85709	Human	sec	AbR85709	Human	sec	762	38	48.7	500	6	ABM23945	Human	sec	AbM23945	Human	sec
690	38	48.7	500	6	ABR99691	Human	sec	AbR99691	Human	sec	763	38	48.7	500	6	ABM05481	Human	sec	AbM05481	Human	sec
691	38	48.7	500	6	ABM00242	Human	sec	AbM00242	Human	sec	764	38	48.7	500	6	ABM15546	Human	sec	AbM15546	Human	sec
692	38	48.7	500	6	ABM00547	Human	sec	AbM00547	Human	sec	765	38	48.7	500	6	ABM08531	Human	sec	AbM08531	Human	sec
693	38	48.7	500	6	ABO29674	Human	sec	AbO29674	Human	sec	766	38	48.7	500	6	ABO42231	Human	sec	AbO42231	Human	sec
694	38	48.7	500	6	ABM23540	Human	sec	AbM23540	Human	sec	767	38	48.7	500	6	ABO37961	Human	sec	AbO37961	Human	sec
695	38	48.7	500	6	ABM29335	Human	sec	AbM29335	Human	sec	768	38	48.7	500	6	ABO45871	Human	sec	AbO45871	Human	sec
696	38	48.7	500	6	ABO38266	Human	sec	AbO38266	Human	sec	769	38	48.7	500	6	ABM66674	Human	sec	AbM66674	Human	sec
697	38	48.7	500	6	ABO45566	Human	PRO	AbO45566	Human	PRO	770	38	48.7	500	6	ADB20182	Human	sec	AdB20182	Human	sec
698	38	48.7	500	6	ABM20490	Human	sec	AbM20490	Human	sec	771	38	48.7	500	6	ABM19575	Human	sec	AbM19575	Human	sec
699	38	48.7	500	6	ADA81341	Human	sec	AdA81341	Human	sec	772	38	48.7	500	6	ABO49287	Human	sec	AbO49287	Human	sec
700	38	48.7	500	6	ABO16604	Human	sec	AbO16604	Human	sec	773	38	48.7	500	6	ABO49592	Human	sec	AbO49592	Human	sec
701	38	48.7	500	6	ABO18230	Human	sec	AbO18230	Human	sec	774	38	48.7	500	6	ADA78434	Human	sec	AdA78434	Human	sec
702	38	48.7	500	6	ABO22857	Human	PRO	AbO22857	Human	PRO	775	38	48.7	500	6	ABO19586	Novel	hum	AbO19586	Novel	hum
703	38	48.7	500	6	ABO22962	Human	sec	AbO22962	Human	sec	776	38	48.7	500	6	ABR88149	Human	sec	AbR88149	Human	sec
704	38	48.7	500	6	ABR92504	Human	sec	AbR92504	Human	sec	777	38	48.7	500	6	ABM26895	Human	sec	AbM26895	Human	sec
705	38	48.7	500	6	ABR81461	Human	sec	AbR81461	Human	sec	778	38	48.7	500	6	ABM03292	Human	sec	AbM03292	Human	sec
706	38	48.7	500	6	ABM77885	Human	sec	AbM77885	Human	sec	779	38	48.7	500	6	ABO33971	Human	sec	AbO33971	Human	sec
707	38	48.7	500	6	ABR89674	Human	sec	AbR89674	Human	sec	780	38	48.7	500	7	ABO49897	Human	sec	AbO49897	Human	sec
708	38	48.7	500	6	ABM26590	Human	sec	AbM26590	Human	sec	781	38	48.7	500	7	ABO50812	Human	sec	AbO50812	Human	sec
709	38	48.7	500	6	ABM13716	Human	sec	AbM13716	Human	sec	782	38	48.7	500	7	ABO05268	Human	sec	AbO05268	Human	sec
710	38	48.7	500	6	ABO28454	Human	sec	AbO28454	Human	sec	783	38	48.7	500	7	ABR74572	Human	sec	AbR74572	Human	sec
711	38	48.7	500	6	ABO30284	Human	sec	AbO30284	Human	sec	784	38	48.7	500	7	ABR77051	Human	sec	AbR77051	Human	sec
712	38	48.7	500	6	ABM07311	Human	sec	AbM07311	Human	sec	785	38	48.7	500	7	ABM17808	Human	sec	AbM17808	Human	sec
713	38	48.7	500	6	ABM03302	Human	sec	AbM03302	Human	sec	786	38	48.7	500	7	ABR95859	Human	sec	AbR95859	Human	sec
714	38	48.7	500	6	ABO37046	Human	sec	AbO37046	Human	sec	787	38	48.7	500	7	ABO21792	Human	sec	AbO21792	Human	sec
715	38	48.7	500	6	ABO41621	Human	sec	AbO41621	Human	sec	788	38	48.7	500	7	ABO19962	Human	sec	AbO19962	Human	sec
716	38	48.7	500	6	ABO35216	Human	PRO	AbO35216	Human	PRO	789	38	48.7	500	7	ABO24265	Human	sec	AbO24265	Human	sec
717	38	48.7	500	6	ABM25065	Human	sec	AbM25065	Human	sec	790	38	48.7	500	7	ABR86014	Human	sec	AbR86014	Human	sec
718	38	48.7	500	6	ABO47457	Human	sec	AbO47457	Human	sec	791	38	48.7	500	7	ABM10666	Human	sec	AbM10666	Human	sec
719	38	48.7	500	6	ABO47762	Human	sec	AbO47762	Human	sec	792	38	48.7	500	7	ABM76665	Human	sec	AbM76665	Human	sec
720	38	48.7	500	6	ABO48372	Human	sec	AbO48372	Human	sec	793	38	48.7	500	7	ABR89369	Human	sec	AbR89369	Human	sec
721	38	48.7	500	6	ABO51422	Human	PRO	AbO51422	Human	PRO	794	38	48.7	500	7	ABM12496	Human	sec	AbM12496	Human	sec
722	38	48.7	500	6	ABO51727	Human	PRO	AbO51727	Human	PRO	795	38	48.7	500	7	ABM05786	Human	sec	AbM05786	Human	sec
723	38	48.7	500	6	ABO50507	Human	sec	AbO50507	Human	sec	796	38	48.7	500	7	ABO34911	Human	PRO	AbO34911	Human	PRO
724	38	48.7	500	6	ABR79631	Human	sec	AbR79631	Human	sec	797	38	48.7	500	7	ABM02987	Human	sec	AbM02987	Human	sec
725	38	48.7	500	6	ABM16893	Human	sec	AbM16893	Human	sec	798	38	48.7	500	7	ABM18965	Human	sec	AbM18965	Human	sec
726	38	48.7	500	6	ABO17925	Human	sec	AbO17925	Human	sec	799	38	48.7	500	7	ABM19270	Human	sec	AbM19270	Human	sec
727	38	48.7	500	6	ABO20877	Human	sec	AbO20877	Human	sec	800	38	48.7	500	7	ABO46481	Human	PRO	AbO46481	Human	PRO
728	38	48.7	500	6	ABR96836	Human	sec	AbR96836	Human	sec	801	38	48.7	500	7	ABO48982	Human	sec	AbO48982	Human	sec
729	38	48.7	500	6	ABM12191	Human	sec	AbM12191	Human	sec	802	38	48.7	500	7	ABR69025	Human	sec	AbR69025	Human	sec
730	38	48.7	500	6	ABM16283	Human	sec	AbM16283	Human	sec	803	38	48.7	500	7	ABR89064	Human	sec	AbR89064	Human	sec
731	38	48.7	500	6	ABM24150	Human	sec	AbM24150	Human	sec	804	38	48.7	500	7	ABR72437	Human	sec	AbR72437	Human	sec
732	38	48.7	500	6	ABM14631	Human	sec	AbM14631	Human	sec	805	38	48.7	500	7	ABR74267	Human	sec	AbR74267	Human	sec
733	38	48.7	500	6	ABM04512	Human	sec	AbM04512	Human	sec	806	38	48.7	500	7	ABO18535	Human	sec	AbO18535	Human	sec
734	38	48.7	500	6	ABM06701	Human	sec	AbM06701	Human	sec	807	38	48.7	500	7	ABR80241	Human	sec	AbR80241	Human	sec
735	38	48.7	500	6	ABM09141	Human	sec	AbM09141	Human	sec	808	38	48.7	500	7	ABM01462	Human	sec	AbM01462	Human	sec
736	38	48.7	500	6	ABO39181	Human	sec	AbO39181	Human	sec	809	38	48.7	500	7	ABM02072	Human	sec	AbM02072	Human	sec
737	38	48.7	500	6	ABO75446	Human	sec	AbO75446	Human	sec	810	38	48.7	500	7	ABR87234	Human	sec	AbR87234	Human	sec
738	38	48.7	500	6	ABM25370	Human	sec	AbM25370	Human	sec	811	38	48.7	500	7	ABM12801	Human	sec	AbM12801	Human	sec
739	38	48.7	500	6	ABM19880	Human	sec	AbM19880	Human	sec	812	38	48.7	500	7	ABM30555	Human	sec	AbM30555	Human	sec
740	38	48.7	500	6	ABO46786	Human	PRO	AbO46786	Human	PRO	813	38	48.7	500	7	ABM24455	Human	sec	AbM24455	Human	sec
741	38	48.7	500	6	ABO47091	Human	PRO	AbO47091	Human	PRO	814	38	48.7	500	7	ABO29369	Human	sec	AbO29369	Human	sec
742	38	48.7	500	6	ADA83139	Human	sec	AdA83139	Human	sec	815	38	48.7	500	7	ABO31199	Human	sec	AbO31199	Human	sec
743	38	48.7	500	6	ABR71522	Human	sec	AbR71522	Human	sec	816	38	48.7	500	7	ABM14326	Human	sec	AbM14326	Human	sec
744	38	48.7	500	6	ABR72132	Human	sec	AbR72132	Human	sec	817	38	48.7	500	7	ABM09751	Human	sec	AbM09751	Human	sec
745	38	48.7	500	6	ABR98471	Human	sec	AbR98471	Human	sec	818	38	48.7	500	7	ABO38876	Human	sec	AbO38876	Human	sec
746	38	48.7	500	6	ABO06841	Human	sec	AbO06841	Human	sec	819	38	48.7	500	7	ABM34641	Human	sec	AbM34641	Human	sec
747	38	48.7	500	6	ABR84794	Human	sec	AbR84794	Human	sec	820	38	48.7	500	7	ABO51117	Human	sec	AbO51117	Human	sec
748	38	48.7	500	6	ABR73352	Human	sec	AbR73352	Human	sec	821	38	48.7	500	7	ABO03943	Human	sec	AbO03943	Human	sec
749	38	48.7	500	6	ABR76446	Human	sec	AbR76446	Human	sec	822	38	48.7	500	7	ABO10413	Human	PRO	AbO10413	Human	PRO
750	38	48.7	500	6	ABR73047	Human	sec	AbR73047	Human	sec	823	38	48.7	500	7	ABR77656	Human	sec	AbR77656	Human	sec
751	38	48.7	500	6	ABM18113	Human	sec	AbM18113	Human	sec											

828	38	48.7	500	7	ABM78190	Human sec	901	38	48.7	500	8	ADE74819	Human sec
829	38	48.7	500	7	ABR89979	Human sec	902	38	48.7	500	8	ADE89798	Human sec
830	38	48.7	500	7	ABM27505	Human sec	903	38	48.7	500	8	ADf61438	Human sec
831	38	48.7	500	7	ABM13106	Human sec	904	38	48.7	500	8	ADf40130	Human sec
832	38	48.7	500	7	ABO31809	Human sec	905	38	48.7	500	8	ADf45926	Human sec
833	38	48.7	500	7	ABM14021	Human sec	906	38	48.7	500	8	ADf24322	Human sec
834	38	48.7	500	7	ABM08226	Human sec	907	38	48.7	500	8	ADf40754	Human sec
835	38	48.7	500	7	ABO40096	Human sec	908	38	48.7	500	8	ADf23698	Human sec
836	38	48.7	500	7	ABM74531	Human sec	909	38	48.7	500	8	ADf33681	Human sec
837	38	48.7	500	7	ABM33726	Human sec	910	38	48.7	500	8	ADf27148	Human sec
838	38	48.7	500	7	ABM20185	Human sec	911	38	48.7	500	8	ADf27784	Human sec
839	38	48.7	500	7	ABO48677	Human sec	912	38	48.7	500	8	ADf41378	Human sec
840	38	48.7	500	7	ABR72742	Human sec	913	38	48.7	500	8	ADf33057	Human sec
841	38	48.7	500	7	ABO15384	Human sec	914	38	48.7	500	8	ADf25423	Human sec
842	38	48.7	500	7	ABR85099	Human sec	915	38	48.7	500	8	ADf26524	Human sec
843	38	48.7	500	7	ABO15079	Human sec	916	38	48.7	500	8	ADf34313	Human sec
844	38	48.7	500	7	ABO17214	Human sec	917	38	48.7	500	8	ADf46550	Human sec
845	38	48.7	500	7	ABM17503	Human sec	918	38	48.7	500	8	ADf96032	Novel hum
846	38	48.7	500	7	ABR85404	Human sec	919	38	48.7	500	8	ADG04303	Novel hum
847	38	48.7	500	7	ABM76970	Human sec	920	38	48.7	500	8	ADG82719	Human PRO
848	38	48.7	500	7	ABO28149	Human sec	921	38	48.7	500	8	ADH26000	Novel hum
849	38	48.7	500	7	ABM22930	Human sec	922	38	48.7	500	8	ADG50536	Human sec
850	38	48.7	500	7	ABM30250	Human sec	923	38	48.7	500	8	ADG49912	Human sec
851	38	48.7	500	7	ABM21710	Human sec	924	38	48.7	500	8	ADG51784	Human sec
852	38	48.7	500	7	ABM21405	Human sec	925	38	48.7	500	8	ADH32969	Human PRO
853	38	48.7	500	7	ABM14936	Human sec	926	38	48.7	500	8	ADG48664	Human sec
854	38	48.7	500	7	ABO41011	Human sec	927	38	48.7	500	8	ADG51160	Human sec
855	38	48.7	500	7	ABO36741	Human sec	928	38	48.7	500	8	ADG48664	Human sec
856	38	48.7	500	7	ABO37351	Human sec	929	38	48.7	500	8	ADG51160	Human sec
857	38	48.7	500	7	ABM75141	Human sec	930	38	48.7	500	8	ADG59104	Human sec
858	38	48.7	500	7	ABM33421	Human sec	931	38	48.7	500	8	ADG62560	Human sec
859	38	48.7	500	7	ABO46176	Human PRO	932	38	48.7	500	8	ADH25585	Human neu
860	38	48.7	500	7	ADH82505	Human sec	933	38	48.7	500	8	ADJ54708	Human PRO
861	38	48.7	500	7	ABM311775	Human sec	934	38	48.7	500	8	ADJ64479	Human PRO
862	38	48.7	500	7	ABM31165	Human sec	935	38	48.7	500	8	ADM31375	Novel hum
863	38	48.7	500	7	ADb73869	Human PRO	936	38	48.7	500	8	ADM17362	Novel hum
864	38	48.7	500	7	ADb85813	Human sec	937	38	48.7	500	8	ADM36422	Novel hum
865	38	48.7	500	7	ABM32080	Human sec	938	38	48.7	500	8	ADM40227	Novel hum
866	38	48.7	500	7	ABM32385	Human sec	939	38	48.7	500	8	ADL07196	Human sec
867	38	48.7	500	7	ABM31470	Human sec	940	38	48.7	500	8	ADN37835	Novel hum
868	38	48.7	500	7	ABM30860	Human sec	941	38	48.7	500	8	ADN98111	Protein f
869	38	48.7	500	7	ADb76585	Human PRO	942	38	48.7	500	8	ADU50318	PRO541, S
870	38	48.7	500	7	ADc44011	Human sec	943	38	48.7	500	9	ADW49597	Human sec
871	38	48.7	500	7	ADc61771	Human sec	944	38	48.7	500	9	ADZ52258	Human sec
872	38	48.7	500	7	ADc63735	Human sec	945	38	48.7	500	3	AAg47662	Arabidops
873	38	48.7	500	7	ADc66835	Human sec	946	38	48.7	500	3	AAg47662	Bacterial
874	38	48.7	500	7	ADc68959	Human sec	947	38	48.7	506	8	ADS44990	Bacterial
875	38	48.7	500	7	ADc63019	Human sec	948	38	48.7	522	4	AAm41693	Human pol
876	38	48.7	500	7	ADc68084	Human sec	949	38	48.7	522	4	AAm41693	Protein f
877	38	48.7	500	7	ADc41404	Human sec	950	38	48.7	522	6	ABR83634	Saccharom
878	38	48.7	500	7	ADc67459	Human sec	951	38	48.7	555	6	ABR83634	Arabidops
879	38	48.7	500	7	ADc62395	Human sec	952	38	48.7	604	4	AAU16970	Human nov
880	38	48.7	500	7	ADc42028	Human sec	953	38	48.7	631	3	AAg29253	Arabidops
881	38	48.7	500	7	ADd05543	Human sec	954	38	48.7	631	3	AAg29253	Arabidops
882	38	48.7	500	7	ADb49397	Human sec	955	38	48.7	631	8	ADT55851	Plant pol
883	38	48.7	500	7	ADb35451	Human sec	956	38	48.7	642	8	ADN17690	Bacterial
884	38	48.7	500	7	ADb16565	Human sec	957	38	48.7	642	8	ADN17690	Bacterial
885	38	48.7	500	7	ADb73180	Human sec	958	38	48.7	648	8	ADf06702	Bacterial
886	38	48.7	500	7	ADb72538	Human sec	959	38	48.7	648	8	ADf06702	Bacterial
887	38	48.7	500	7	ADf17189	Human sec	960	38	48.7	642	8	ADf06702	Bacterial
888	38	48.7	500	7	ADf47203	Human sec	961	38	48.7	642	8	ADf06702	Bacterial
889	38	48.7	500	7	ADG02538	Novel hum	962	38	48.7	642	8	ADf06702	Bacterial
890	38	48.7	500	7	ADG01245	Novel hum	963	38	48.7	642	8	ADf06702	Bacterial
891	38	48.7	500	7	ADG95420	Novel hum	964	38	48.7	642	8	ADf06702	Bacterial
892	38	48.7	500	7	ADH12235	Novel hum	965	38	48.7	642	8	ADf06702	Bacterial
893	38	48.7	500	7	ADH08895	Human PRO	966	38	48.7	642	8	ADf06702	Bacterial
894	38	48.7	500	7	ADG52960	Human sec	967	38	48.7	642	8	ADf06702	Bacterial
895	38	48.7	500	7	ADG60280	Human sec	968	38	48.7	642	8	ADf06702	Bacterial
896	38	48.7	500	7	ADf161040	Human sec	969	38	48.7	642	8	ADf06702	Bacterial
897	38	48.7	500	7	ADf32676	Novel hum	970	38	48.7	642	8	ADf06702	Bacterial
898	38	48.7	500	7	ADM30210	Novel hum	971	37.5	48.1	14	2	AAW09335	Thrombopo
899	38	48.7	500	8	ADe48697	Human sec	972	37.5	48.1	14	2	AAW09335	Thrombopo
900	38	48.7	500	8	ADe74207	Human, sec	973	37.5	48.1	14	2	AAW46891	Cyclic pe

974 37.5 48.1 14 3 AAB16997 TPO-mimet
 975 37.5 48.1 14 4 AAU25905 Human thr
 976 37.5 48.1 14 5 ABB72883 TPO minet
 977 37.5 48.1 14 6 ABG71743 TPO recep
 978 37.5 48.1 14 7 ADJ73034 TPO minet
 979 37.5 48.1 14 8 ADJ52669 CH1 delet
 980 37.5 48.1 14 8 ADJ51630 CH1 delet
 981 37 47.4 9 4 AAM22332 HIV pepti
 982 37 47.4 9 4 AAM22645 HIV pepti
 983 37 47.4 9 4 AAM22647 HIV pepti
 984 37 47.4 9 4 AAM22646 HIV pepti
 985 37 47.4 10 2 AAU09560 Thrombopo
 986 37 47.4 10 2 AAW36711 Thrombopo
 987 37 47.4 10 4 AAU25930 Human thr
 988 37 47.4 12 2 AAW09574 Thrombopo
 989 37 47.4 12 2 AAW09579 Thrombopo
 990 37 47.4 12 2 AAW36725 Thrombopo
 991 37 47.4 12 2 AAW36730 Thrombopo
 992 37 47.4 12 4 AAU25949 Human thr
 993 37 47.4 12 4 AAU25944 Human thr
 994 37 47.4 15 8 ADN17047 Gi-alpha
 995 37 47.4 17 2 AAW09607 Thrombopo
 996 37 47.4 17 2 AAW09600 Thrombopo
 997 37 47.4 17 2 AAW09610 Thrombopo
 998 37 47.4 17 2 AAW36761 Thrombopo
 999 37 47.4 17 2 AAW36758 Thrombopo
 1000 37 47.4 17 2 AAW36751 Thrombopo

ALIGNMENTS

RESULT 1
 AAW09524
 ID AAW09524 standard; protein; 12 AA.

XX AC AAW09524;

DT 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion, chemotherapy; radiation therapy.

XX Synthetic.

XX WO9640189-A1.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.
 XX thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 28; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. NO. 6.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12

Db 1 CTLREWLHGGFC 12

RESULT 2

AAW36675
 ID AAW36675 standard; peptide; 12 AA.

XX AC AAW36675;

XX 27-FEB-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX WO9640750-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009623.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 28; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. NO. 6.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
 Db 1 CTLREWLHGGFC 12

RESULT 3
 AAB16993
 ID AAB16993 standard; peptide; 12 AA.
 XX AAB16993;
 AC AAB16993;
 DT 31-OCT-2000 (first entry)
 XX TPO-mimetic peptide sequence SEQ ID NO:49.
 DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX Synthetic.
 OS WO200024782-A2.
 PN 04-MAY-2000.
 XX 25-OCT-1999; 99WO-US025044.
 XX 23-OCT-1998; 98US-0105371P.
 PR 22-OCT-1999; 99US-00428082.
 XX (AMGE-) AMGEN INC.
 PA Feige U, Liu C, Cheetham J, Boone TC;
 PI WPI; 2000-350702/30.
 DR Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.
 XX Claim 19; Page 212; 608pp; English.

The present invention describes composition of matter (I) comprising an
 Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 P3, and P4 = are each independently sequences of pharmacologically active
 peptides; L1, L2, L3, and L4 = are each independently linker; and a, b,
 c, d, e, and f = are each independently 0 or 1, provided that at least 1
 of a and b is 1. The composition can have cytostatic, antiasthmatic,
 thrombolytic and immunosuppressive activities. DNAs, vectors and host
 cells from the present invention can be used for producing pharmaceutical
 compositions. The compositions are useful for treating cancer, asthma,
 thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 a Fab domain) can provide a longer half-life or incorporate functions
 such as Fc receptor binding, protein A binding, complement fixation, and
 possibly placental transfer. AAB69443 to AAB69526 and AAB16955 to
 AAB18003 represent nucleotide and amino acid sequences used in the
 exemplification of the present invention

Sequence 12 AA;
 Query Match 100.0%; Score 78; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
 Db 1 CTLREWLHGGFC 12

RESULT 4
 AAU25894
 ID AAU25894 standard; peptide; 12 AA.
 XX AAU25894;
 AC AAU25894;
 DT 17-DEC-2001 (first entry)
 XX Human thrombopoietin receptor (TPO-R) activator peptide #80.
 DE Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX Homo sapiens.
 OS US6251864-B1.
 PN 26-JUN-2001.
 XX 01-MAR-2000; 2000US-00516704.
 XX 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US009623.
 PR 15-AUG-1996; 96US-00699027.
 XX (GLAXO) GLAXO GROUP LTD.
 PA Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 DR Activating thrombopoietin receptors in cells, used to treat
 XX thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX Disclosure; Col 21; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 of activating thrombopoietin receptors in cells comprise contacting the
 cells with effective amounts of peptides and peptide mimetics attached to
 hydrophilic polymers. The methods are used to treat thrombocytopenia such
 as that due to chemotherapy, radiation therapy or bone-marrow
 transplantation and to prevent thrombocytopenia in patients at risk. The
 sequences are used to treat and prevent hematological disorders
 including thrombocytopenia and platelet disorders. They are used in vitro
 as unique tools for understanding the biological role of thrombopoietin
 (TPO) and to develop other compounds that bind to and activate the TPO
 receptor. The peptides can be used to detect TPO receptors on living
 cells and fixed cells, in biological fluids, in tissue homogenates, and
 in purified or natural biological materials. They may also be used for in
 situ staining, fluorescence-activated cell sorting, Western blotting and
 enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 be used for in vitro expansion of megakaryocytes and their committed
 progenitors alone or in conjunction with additional cytokines

Sequence 12 AA;
 Query Match 100.0%; Score 78; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
 Db 1 CTLREWLHGGFC 12

Db 1 CTLREWLHGGFC 12

RESULT 5

ABB72879

ID ABB72879 standard; peptide; 12 AA.

XX

AC ABB72879;

XX

DT 05-APR-2002 (first entry)

XX

DE TPO mimetic peptide SEQ ID NO:49.

XX

XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor; TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP; TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist; MMP inhibitor; antiinflammatory; antitumour; immunosuppressive; cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological; antianemic; anorectic; antiinfertility; haemostatic; dermatological; neuroprotective; inflammatory disease; autoimmune disease; tumour growth; cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity; sleep disorder; neurological degenerative disease; anaemia; thrombocytopaenia; metastatic tumour; systemic lupus erythematosus; Fanconi's syndrome.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200183525-A2.

XX

PD 08-NOV-2001.

XX

PF 02-MAY-2001; 2001WO-US014310.

XX

PP 03-MAY-2000; 2000US-00563286.

XX

PR (AMGE-) AMGEN INC.

XX

PA Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

PI WPI, 2002-130313/17.

XX

DR Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

XX

PS Claim 39; Page 43; 176pp; English.

XX

CC The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cycostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-inertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05; Mismatches 0; Gaps 0;
Matches 12; Conservative 0; Indels 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | | | | | | | |
Db 1 CTLREWLHGGFC 12

RESULT 6

ADJ73030

ID ADJ73030 standard; peptide; 12 AA.

XX

AC ADJ73030;

XX

DT 06-MAY-2004 (first entry)

XX

DE TPO mimetic peptide sequence SeqID 484.

XX

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune; cardiovascular; infectious; malignant; neurologic disease; anaemia; immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective; TPO.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

PI WPI, 2003-804237/75.

XX

DR New CDR mimetibody comprising a portion of a heavy or light chain variable region comprising human framework or ligand binding region, useful for preparing a composition for treating e.g., immune, cardiovascular or neurologic disease.

XX

PS Disclosure; SEQ ID NO 484; 97pp; English.

XX

CC This invention relates to novel mammalian CDR mimetibodies, specific portions or variants thereof. Specifically, it refers to an antibody fragment where a protein has been inserted into, or replaces a portion of, one or more CDR regions, such that each CDR mimetibody comprises at least one portion of a heavy chain or light chain variable region, which itself comprises at least one human framework region and at least one ligand binding region (LBR). The present invention describes human mimetibodies, including modified immunoglobulins and cleavage products that can be useful in gene therapy and the generation of transgenic plants and animals. Furthermore, the CDR mimetibody is useful for preparing compositions for modulating, treating or reducing the symptoms of immune, cardiovascular, infectious, malignant and/or neurologic diseases, as well as anaemia. Accordingly, they exhibit immunomodulator, cardiant, antimicrobial, cytostatic and neuroprotective activities. This peptide sequence is a TPO mimetic peptide sequence used to make a mimetibody of the invention.

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 7; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05; Mismatches 0; Gaps 0;
Matches 12; Conservative 0; Indels 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | | | | | | | |

```

Db          1 CTLREWLHGGFC 12

RESULT 7
ADJ52665
ID ADJ52665 standard; peptide; 12 AA.
XX
XX
AC ADJ52665;
XX
XX 06-MAY-2004 (first entry)
XX
XX CH1 deleted mimetibody-related peptide SeqID484.
XX
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
XX Unidentified.
OS Synthetic.
XX
XX WO2004002417-A2.
XX
XX 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Neseppor TC;
PI Kucoloski KA;
XX
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
XX Claim 2; SEQ ID NO 484; 129pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
XX Sequence 12 AA;
XX
Query Match 100.0%; Score 78; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
    |||||
Db 1 CTLREWLHGGFC 12

```

```

RESULT 8
ADJ51626
ID ADJ51626 standard; peptide; 12 AA.
XX
XX ADJ51626;
XX
XX 06-MAY-2004 (first entry)
XX
XX CH1 deleted mimetibody-related peptide SeqID484.
XX
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
XX Unidentified.
OS Synthetic.
XX
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Neseppor TC;
PI Kucoloski KA;
XX
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 14; SEQ ID NO 484; 123pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
XX Sequence 12 AA;
XX

```

Query Match 100.0%; Score 78; DB 8; Length 12;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
 |||||
 Db 1 CTLREWLHGGFC 12

RESULT 9

AAW09464
 ID AAW09464 standard; protein; 16 AA.

AC AAW09464;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding compound peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

XX WO9640189-A1.

PD 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 DR WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of haematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

PS Claim 23; Page 90; 106pp; English.

XX The present sequence is a compound which binds to thrombopoietin (TPO)
 CC receptor (TR). It has a molecular weight of < 8000 Da, and a binding
 CC affinity to TR as expressed by an IC50 of no more than about 100 nm. The
 CC compound can be used for treating patients suffering from haematological
 CC disorders and thrombocytopenia resulting from chemotherapy, radiation
 CC therapy or bone marrow transfusions. The peptide may also be used to
 CC maintain the proliferation and growth of TPO-dependent cell lines and for
 CC use in biological research, for detecting TPO receptors on living cells.
 CC The present sequence is a specifically claimed example of a compound that
 CC is administered to a patient suffering from a disorder that is
 CC susceptible to treatment with a thrombopoietin agonist

XX Sequence 16 AA;

Query Match 100.0%; Score 78; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
 |||||
 Db 3 CTLREWLHGGFC 14

RESULT 10

AAW33329
 ID AAW33329 standard; peptide; 16 AA.

XX AAW33329;

DT 11-MAR-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

OS Synthetic.

XX WO9640750-A1.

PD 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

DR WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Claim 14; Page 89; 106pp; English.

XX The present peptide binds the thrombopoietin receptor (TR), has a
 CC molecular weight of less than 8000 Da and a TR binding affinity as
 CC expressed by an IC50 of no more than about 100 microm. It can be used to
 CC treat disorders which are susceptible to treatment with a thrombopoietin
 CC agonist, preferably haematological disorders and thrombocytopenia
 CC resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. It can also be used diagnostically, e.g. to investigate the
 CC mechanism of thrombopoietin signal transduction and receptor activation,
 CC or to maintain the proliferation and growth of thrombopoietin dependent
 CC cell lines

XX Sequence 16 AA;

Query Match 100.0%; Score 78; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
 |||||
 Db 3 CTLREWLHGGFC 14

RESULT 11

AAW17019
 ID AAW17019 standard; peptide; 16 AA.

XX AAW17019;

DT 31-OCT-2000 (first entry)

XX TPO-mimetic peptide sequence SEQ ID NO:75.

DE Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;

immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 vascular endothelial growth factor; matrix metalloproteinase; asthma;
 thrombosis; pharmaceutical.
 Synthetic.
 WO200024782-A2.
 04-MAY-2000.
 25-OCT-1999; 99WO-US025044.
 23-OCT-1998; 98US-0105371P.
 22-OCT-1999; 99US-00428082.
 (AMGE-) AMGEN INC.
 Feige U, Liu C, Cheetham J, Boone TC;
 WPI; 2000-350702/30.
 Novel composition of matter comprising an Fc domain and pharmacologically
 active peptides, useful for treating cancer and autoimmune diseases.
 Claim 19; Page 220; 608pp; English.
 The present invention describes composition of matter (I) comprising an
 Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 P3, and P4 = are each independently sequences of pharmacologically active
 peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 c, d, e, and f = are each independently 0 or 1, provided that at least 1
 of a and b is 1. The composition can have cytostatic, antiasthmatic,
 thrombolytic and immunosuppressive activities. DNAs, vectors and host
 cells from the present invention can be used for producing pharmaceutical
 compositions. The compositions are useful for treating cancer, asthma,
 thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 a Fab domain) can provide a longer half-life or incorporate functions
 such as Fc receptor binding, protein A binding, complement fixation, and
 possibly placental transfer. AA69443 to AA69526 and AAB16955 to
 AAB18003 represent nucleotide and amino acid sequences used in the
 exemplification of the present invention
 Sequence 16 AA;
 Query Match 100.0%; Score 78; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTLREWLHGSGFC 12
 |||||
 Db 3 CTLREWLHGSGFC 14
 RESULT 12
 AAU25829
 ID AAU25829 standard; peptide; 16 AA.
 XX
 AC AAU25829;
 XX
 DT 17-DEC-2001 (first entry)
 XX
 DE Human thrombopoietin receptor (TPO-R) activator peptide #15.
 XX
 KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;

in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 Homo sapiens.
 US6251864-B1.
 26-JUN-2001.
 01-MAR-2000; 2000US-00516704.
 07-JUN-1995; 95US-00478128.
 07-JUN-1995; 95US-00485301.
 07-JUN-1996; 96WO-US009823.
 15-AUG-1996; 96US-00699027.
 (GLAXO) GLAXO GROUP LTD.
 Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
 Yin Q;
 WPI; 2001-564142/63.
 Activating thrombopoietin receptors in cells, used to treat
 thrombocytopenia and hematological disorders, comprises contacting cells
 with peptides and peptide mimetics attached to hydrophilic polymers.
 Disclosure; Col 69-70; 128pp; English.
 Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 of activating thrombopoietin receptors in cells comprise contacting the
 cells with effective amounts of peptides and peptide mimetics attached to
 hydrophilic polymers. The methods are used to treat thrombocytopenia such
 as that due to chemotherapy, radiation therapy or bone-marrow
 transplantation and to prevent thrombocytopenia in patients at risk. The
 sequences are used to treat and prevent hematological disorders
 including thrombocytopenia and platelet disorders. They are used in vitro
 as unique tools for understanding the biological role of thrombopoietin
 (TPO) and to develop other compounds that bind to and activate the TPO
 receptor. The peptides can be used to detect TPO receptors on living
 cells and fixed cells, in biological fluids, in tissue homogenates, and
 in purified or natural biological materials. They may also be used for in
 situ staining, fluorescence-activated cell sorting, Western blotting and
 enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 be used for in vitro expansion of megakaryocytes and their committed
 progenitors alone or in conjunction with additional cytokines
 Sequence 16 AA;
 Query Match 100.0%; Score 78; DB 4; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTLREWLHGSGFC 12
 |||||
 Db 3 CTLREWLHGSGFC 14
 RESULT 13
 ABB72905
 ID ABB72905 standard; peptide; 16 AA.
 XX
 AC ABB72905;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE TPO mimetic peptide SEQ ID NO:75.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antinaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200183525-A2.
XX
XX PD 08-NOV-2001.
XX
XX PF 02-MAY-2001; 2001WO-US014310.
XX
XX PR 03-MAY-2000; 2000US-00563286.
XX
XX PA (AMGE-) AMGEN INC.
XX
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
XX DR WPI; 2002-130313/17.
XX
XX PT Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX diabetic retinopathy, obesity, sleep disorders and infertility.
XX
XX PS Claim 39; Page 44; 176pp; English.
XX
XX CC The present invention describes a vehicle-peptide molecule (I) or its
XX multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
XX cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
XX antinaemic, anorectic, antiinfertility, haemostatic, dermatological and
XX neuroprotective activities. (I) can be used as a therapeutic or
XX prophylactic agent as well as for screening purposes. (I) is useful for
XX diagnosing diseases characterised by dysfunction of their associated
XX protein of interest, for identifying normal or abnormal proteins of
XX interest, as a part of diagnostic kit to detect the presence of their
XX proteins of interest in a biological sample. Additionally, (I) is useful
XX for treating inflammatory and autoimmune diseases, tumour growth, cancer,
XX rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
XX infertility, and neurological degenerative diseases. (I), comprising EPO-
XX mimetic compounds are useful for treating disorders characterised by low
XX red blood cell levels such as anaemia. The TPO-mimetic comprising
XX compounds are useful for treating conditions that involve an existing
XX megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
XX deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
XX tumour which result in thrombocytopenia, systemic lupus erythematosus,
XX and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
XX represent amino acid and nucleic acid sequences used in the
XX exemplification of the present invention

SQ Sequence 16 AA;
Query Match 100.0%; Score 78; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | | | |
Db 3 CTLREWLHGGFC 14

RESULT 14
ADJ73057
ID ADJ73057 standard; peptide; 16 AA.
XX
XX AC ADJ73057;
XX
XX DT 06-MAY-2004 (first entry)

XX TPO mimetic peptide sequence SeqID 511.
DE
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
XX PN WO2003084477-A2.
XX
XX PD 16-OCT-2003.
XX
XX PF 24-MAR-2003; 2003WO-US009139.
XX
XX PR 29-MAR-2002; 2002US-0368791P.
XX
XX PA (CENZ) CENTOCOR INC.
XX
XX PI Heavner GA, Knight DM, Scallan BJ, Ghayeb J;
XX
XX DR WPI; 2003-804237/75.
XX
XX PT New CDR mimetibody comprising a portion of a heavy or light chain
XX variable region comprising human framework or ligand binding region,
XX useful for preparing a composition for treating e.g., immune,
XX cardiovascular or neurologic disease.
XX
XX PS Disclosure; SEQ ID NO 511; 97pp; English.
XX
XX CC This invention relates to novel mammalian CDR mimetibodies, specific
XX portions or variants thereof. Specifically, it refers to an antibody
XX fragment where a protein has been inserted into, or replaces a portion
XX of, one or more CDR regions, such that each CDR mimetibody comprises at
XX least one portion of a heavy chain or light chain variable region, which
XX itself comprises at least one human framework region and at least one
XX ligand binding region (LBR). The present invention describes human
XX mimetibodies, including modified immunoglobulins and cleavage products
XX that can be useful in gene therapy and the generation of transgenic
XX plants and animals. Furthermore, the CDR mimetibody is useful for
XX preparing compositions for modulating, treating or reducing the symptoms
XX of immune, cardiovascular, infectious, malignant and/or neurologic
XX diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
XX cardiant, antimicrobial, cytostatic and neuroprotective activities. This
XX peptide sequence is a TPO mimetic peptide sequence used to make a
XX mimetibody of the invention.

SQ Sequence 16 AA;

Query Match 100.0%; Score 78; DB 7; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | | | |
Db 3 CTLREWLHGGFC 14

RESULT 15
ADJ52692
ID ADJ52692 standard; peptide; 16 AA.
XX
XX AC ADJ52692;
XX
XX DT 06-MAY-2004 (first entry)

XX CH1 deleted mimetibody-related peptide SeqID511.
DE
XX
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX Unidentified.
 OS Synthetic.
 XX WO2004002417-A2.
 FN 08-JAN-2004.
 PD 27-JUN-2003; 2003WO-US020347.
 XX 28-JUN-2002; 2002US-0392431P.
 PR (CENZ) CENTOCOR INC.
 XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutoloski KA;
 XX WPI; 2004-082870/08.
 DR New CHI deleted mimetibody polypeptides and nucleic acids, useful for
 XX modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 PS Claim 2; SEQ ID NO 511; 129pp; English.
 XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, neurotropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 78; DB 8; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTLREWLHGFC 12
 DB 3 CTLREWLHGFC 14
 RESULT 16
 ID ADJ51653
 XX ADJ51653 standard; peptide; 16 AA.
 AC ADJ51653;
 XX 06-MAY-2004 (first entry)
 DT CHI deleted mimetibody-related peptide SeqID511.
 XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.
 XX Unidentified.
 OS Synthetic.
 XX WO2004002424-A2.
 FN 08-JAN-2004.
 PD 30-JUN-2003; 2003WO-US020495.
 XX 28-JUN-2002; 2002US-0392431P.
 PR 19-SEP-2002; 2002US-0412144P.
 XX (CENZ) CENTOCOR INC.
 PA Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 XX Kutoloski KA;
 PI WPI; 2004-082872/08.
 DR New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 XX diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.
 XX Claim 15; SEQ ID NO 511; 123pp; English.
 PS This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
 CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.
 XX Sequence 16 AA;
 SQ
 Query Match 100.0%; Score 78; DB 8; Length 16;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTLREWLHGFC 12
 DB 3 CTLREWLHGFC 14
 RESULT 17
 ID AAW09582
 XX AAW09582 standard; protein; 12 AA.
 AC AAW09582;
 XX 10-SEP-1997 (first entry)
 DT
 XX

Thrombopoietin receptor binding peptide.

DE XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX OS
XX Synthetic.
XX
XX WO9640189-A1.
XX
XX PD 19-DEC-1996.
XX
XX PF 05-JUN-1996; 96WO-US008998.
XX
XX PR 07-JUN-1995; 95US-00472371.
XX PR 07-JUN-1995; 95US-00473604.
XX PR 07-JUN-1995; 95US-00476168.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00484090.
XX PR 07-JUN-1995; 95US-00485301.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 29; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
XX Sequence 12 AA;
XX
XX Query Match 75.6%; Score 59; DB 2; Length 12;
XX Best Local Similarity 75.0%; Pred. No. 0.05;
XX Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX Qy 1 CTLREWLHGGFC 12
XX |||:|||||
XX Db 1 CTLQEWLSGGGLC 12
XX
XX RESULT 18
XX AAW36733
XX ID AAW36733 standard; peptide; 12 AA.
XX
XX AC AAW36733;
XX
XX DT 09-MAR-1998 (first entry)
XX
XX DE Thrombopoietin receptor binding peptide.
XX
XX Thrombopoietin receptor; binding peptide; treatment; agonist;
XX haematological disorder; thrombocytopenia; chemotherapy;
XX radiation therapy; bone marrow transfusion; diagnosis;
XX signal transduction; receptor activation; cell culture.
XX
XX OS Synthetic.
XX
XX WO9640750-A1.
XX
XX PD 19-DEC-1996.
XX

PF 07-JUN-1996; 96WO-US009623.
XX
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX
XX PA (GLAX) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX PT Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX PS Disclosure; Page 29; 106pp; English.
XX
XX CC The present peptide, which binds the thrombopoietin receptor (TR), can be
XX used to treat disorders which are susceptible to treatment with a
XX thrombopoietin agonist, preferably haematological disorders and
XX thrombocytopenia resulting from chemotherapy, radiation therapy or bone
XX marrow transfusions. It can also be used diagnostically, e.g. to
XX investigate the mechanism of thrombopoietin signal transduction and
XX receptor activation, or to maintain the proliferation and growth of
XX thrombopoietin dependent cell lines
XX
XX SQ Sequence 12 AA;
XX
XX Query Match 75.6%; Score 59; DB 2; Length 12;
XX Best Local Similarity 75.0%; Pred. No. 0.05;
XX Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX Qy 1 CTLREWLHGGFC 12
XX |||:|||||
XX Db 1 CTLQEWLSGGGLC 12
XX
XX RESULT 19
XX AAU25952
XX ID AAU25952 standard; peptide; 12 AA.
XX
XX AC AAU25952;
XX
XX DT 17-DEC-2001 (first entry)
XX
XX DE Human thrombopoietin receptor (TPO-R) activator peptide #138.
XX
XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
XX bone marrow transplantation; haematological disorder; platelet disorder;
XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;
XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;
XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
XX OS Homo sapiens.
XX
XX PN US6251864-B1.
XX
XX PD 26-JUN-2001.
XX
XX PF 01-MAR-2000; 2000US-00516704.
XX
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00485301.
XX PR 07-JUN-1996; 96WO-US009623.
XX PR 15-AUG-1996; 96US-00699027.
XX
XX PA (GLAX) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
XX Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S;
XX Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

XX thrombocytopenia and hematological disorders, comprises contacting cells

PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 22; 128pp; English.

XX Sequences AAU25015-AAU26049 represent peptides and peptide mimetics that

CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods

CC of activating thrombopoietin receptors in cells comprise contacting the

CC cells with effective amounts of peptides and peptide mimetics attached to

CC hydrophilic polymers. The methods are used to treat thrombocytopenia such

CC as that due to chemotherapy, radiation therapy or bone-marrow

CC transplantation and to prevent thrombocytopenia in patients at risk. The

CC sequences are used to treat and prevent hematological disorders

CC including thrombocytopenia and platelet disorders. They are used in vitro

CC as unique tools for understanding the biological role of thrombopoietin

CC (TPO) and to develop other compounds that bind to and activate the TPO

CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and

CC in purified or natural biological materials. They may also be used for in

CC situ staining, fluorescence-activated cell sorting, Western blotting and

CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

CC be used for in vitro expansion of megakaryocytes and their committed

CC progenitors alone or in conjunction with additional cytokines

XX Sequence 12 AA;

SQ Query Match 75.6%; Score 59; DB 4; Length 12;

Best Local Similarity 75.0%; Pred. NO. 0.05; 2; Indels 0; Gaps 0;

Matches 9; Conservative 1; Mismatches 2;

Qy 1 CTLEWLHGGFC 12

Db 1 CTLEWLHGGFC 12

RESULT 20

AAW09572

ID AAW09572 standard; protein; 12 AA.

XX AC AAW09572;

XX 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;

KW bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX WO9640189-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.

XX 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide

PT mimetic(s) - useful in treatment of hematological disorders, esp.

PT thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 29; 106pp; English.

XX The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering

CC from hematological disorders and thrombocytopenia resulting from

CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide

CC may also be used to maintain the proliferation and growth of TPO-

CC dependent cell lines and for use in biological research, for detecting

CC TPO receptors on living cells

XX Sequence 12 AA;

SQ Query Match 69.2%; Score 54; DB 2; Length 12;

Best Local Similarity 66.7%; Pred. NO. 0.28; 3; Indels 0; Gaps 0;

Matches 8; Conservative 1; Mismatches 3;

Qy 1 CTLEWLHGGFC 12

Db 1 CTLEWLHGGFC 12

RESULT 21

AAW36723

ID AAW36723 standard; peptide; 12 AA.

XX AC AAW36723;

XX 09-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW hematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

XX Synthetic.

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the

PT thrombopoietin receptor - useful in treatment of hematological

PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 29; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

CC used to treat disorders which are susceptible to treatment with a

CC thrombopoietin agonist, preferably hematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone

CC marrow transfusions. It can also be used diagnostically, e.g. to

CC investigate the mechanism of thrombopoietin signal transduction and

CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 2; Length 12;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
 |||||: ||
 Db 1 CTLREWVFAGLC 12

RESULT 22

AA16994
 ID AAB16994 standard; peptide; 12 AA.
 XX AC AAB16994;
 XX DT 31-OCT-2000 (first entry)
 XX DE TPO-mimetic peptide sequence SEQ ID NO:50.
 XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX OS Synthetic.
 XX PN WO200024782-A2.
 XX PD 04-MAY-2000.
 XX PF 25-OCT-1999; 99WO-US025044.
 XX PR 23-OCT-1998; 98US-0105371P.
 XX PR 22-OCT-1999; 99US-00428082.
 XX PA (AMGE-) AMGEN INC.
 XX PI Feige U, Liu C, Cheetham J, Boone TC;
 XX DR WPI; 2000-350702/30.
 XX PT Novel composition of matter comprising an Fc domain and pharmacologically
 active peptides, useful for treating cancer and autoimmune diseases.
 XX PS Claim 19; Page 212; 608pp; English.

CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)-a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAG4943 to AAG69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 3; Length 12;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
 |||||: ||
 Db 1 CTLREWVFAGLC 12

RESULT 23

AAU25942
 ID AAU25942 standard; peptide; 12 AA.
 XX AC AAU25942;
 XX DT 17-DEC-2001 (first entry)
 XX DE Human thrombopoietin receptor (TPO-R) activator peptide #128.
 XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX OS Homo sapiens.
 XX PN US6251864-B1.
 XX PD 26-JUN-2001.
 XX PF 01-MAR-2000; 2000US-00516704.
 XX PR 07-JUN-1995; 95US-00478128.
 XX PR 07-JUN-1995; 95US-00485301.
 XX PR 07-JUN-1996; 96WO-US009623.
 XX PR 15-AUG-1996; 96US-00699027.
 XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Dower WJ, Barrett RW, Cwiria SR, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX DR WPI; 2001-564142/63.
 XX PT Activating thrombopoietin receptors in cells, used to treat
 thrombocytopenia and hematological disorders, comprises contacting cells
 with peptides and peptide mimetics attached to hydrophilic polymers.
 XX PS Disclosure; Col 22; 128pp; English.

CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines

Mon May 15 11:35:35 2006

Query Match 69.2%; Score 54; DB 4; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

QY 1 CTLREWLHGGFC 12
| | | | | : | |
DB 1 CTLREWVFAGLC 12

Query Match 69.2%; Score 54; DB 5; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | : | |
DB 1 CTLREWVFAGLC 12

RESULT 25
ADJ73031.
ID ADJ73031 standard; peptide; 12 AA.

XX AC ADJ73031;
XX DT 06-MAY-2004 (first entry)
XX DE TPO mimetic peptide sequence SeqID 485.
XX KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiac; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX OS Synthetic.
XX PN WO2003084477-A2.
XX PD 16-OCT-2003.
XX PF 24-MAR-2003; 2003WO-US009139.
XX PR 29-MAR-2002; 2002US-0368791P.
XX PA (CENZ) CENTOCOR INC.
XX PI Heavner GA, Knight DM, Scallon BJ, Ghraieb J;
XX DR WPI; 2003-804237/75.
XX PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX PS Disclosure; SEQ ID NO 485; 97pp; English.
XX CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiac, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX SQ Sequence 12 AA;

QY 1 CTLREWLHGGFC 12
| | | | | : | |
DB 1 CTLREWVFAGLC 12

Query Match 69.2%; Score 54; DB 4; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
| | | | | : | |
DB 1 CTLREWVFAGLC 12

RESULT 24
ABB72880
ID ABB72880 standard; peptide; 12 AA.
XX AC ABB72880;
XX DT 05-APR-2002 (first entry)
XX DE TPO mimetic peptide SEQ ID NO:50.
XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TWP;
KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200183525-A2.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-US014310.
XX PR 03-MAY-2000; 2000US-00563286.
XX PA (AMGE-) AMGEN INC.
XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JW;
XX DR WPI; 2002-130313/17.
XX PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX PS Claim 39; Page 43; 176pp; English.
XX CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels, such as anaemia. The TPO-mimetic comprising
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

Query Match 69.2%; Score 54; DB 7; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWVFAGLC 12

RESULT 26

ID ADJ52666 standard; peptide; 12 AA.

XX AC ADJ52666;

XX DT 06-MAY-2004 (first entry)

XX DE CHI deleted mimetibody-related peptide SeqID485.

XX KW CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX OS Unidentified.

XX OS Synthetic.

XX PN WO2004002417-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020347.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

XX PI Kutoloski KA;

XX DR WPI; 2004-082870/08.

XX PT New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
modulating, treating, alleviating, preventing an immune, cardiovascular,
or neurodegenerative disease or disorder, anemia, cancer, or infectious
diseases.

XX PS Claim 2; SEQ ID NO 485; 129pp; English.

XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
which encode them), compositions, methods and uses. The invention may be
useful for the development of compounds with an immunosuppressive,
cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
antibacterial, virucide or fungicide activity. In addition, the disclosed
sequences may prove useful for gene therapy. The CHI-deleted mimetibody
is useful for diagnosing or treating a disease condition in a cell,
tissue, organ or animal, specifically for modulating, treating,
alleviating, preventing the incidence or reducing the symptoms of an
immune, cardiovascular (for example arrhythmia, hypertension or heart
failure), or neurodegenerative (for example multiple sclerosis, dementia
or Alzheimer's disease) diseases or disorders, anaemia, cancerous
conditions, or infectious diseases (for example bacterial, viral or
fungal infection). The present sequence is that of a peptide which may be
used during the creation of a mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.28;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12
| | | | | | | |
Db 1 CTLREWVFAGLC 12

RESULT 27

ID ADJ51627 standard; peptide; 12 AA.

XX AC ADJ51627;

XX DT 06-MAY-2004 (first entry)

XX DE CHI deleted mimetibody-related peptide SeqID485.

XX KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW anti-allergic; muscular-Gen; cytostatic; anti-inflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.

XX OS Unidentified.

XX OS Synthetic.

XX PN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

XX PI Kutoloski KA;

XX DR WPI; 2004-082872/08.

XX PT New CHI deleted mimetibody polypeptide and nucleic acid, useful for
diagnosing, preventing or treating cardiovascular, dermatologic,
endocrine, gastrointestinal, gynecologic, infectious, neurologic and
nutritional disorders.

XX PS Claim 14; SEQ ID NO 485; 123pp; English.

XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
which encode them), compositions, methods and uses. The invention may be
useful for the development of compounds with an osteopathic,
cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
immunomodulator, anti-allergic, muscular-Gen, cytostatic,
anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
modulator or cytokine-agonist. The methods and compositions of the
present invention are useful for the diagnosis, prevention and/or
treatment of diseases or conditions associated with aberrant expression
or activity of the CHI deleted mimetibody, such as a bone or joint,
cardiovascular, dental or oral, dermatological, ear, nose or throat,
endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
obstetric, haematologic, immunological, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX SQ Sequence 12 AA;

Query Match 69.2%; Score 54; DB 8; Length 12;
 Best Local Similarity 66.7%; Pred. No. 0.28;
 Matches 8; Conservative 1; Mismatches 0; Indels 3; Gaps 0;

Qy 1 CTLREWLHGGFC 12

Db 1 CTLREWVFLG 12

RESULT 28
 AAW09581
 ID AAW09581 standard; protein; 12 AA.

XX AC AAW09581;

XX DT 10-SEP-1997 (first entry)

XX DE Thrombopoietin receptor binding peptide.

XX KW Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.

XX OS Synthetic.

XX PN WO9640189-A1.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PR 07-JUN-1995; 95US-00485301.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwirle SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX DR WPI; 1997-051883/05.

XX PT Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.
 PT thrombocytopenia resulting from chemotherapy, etc.

XX PS Disclosure; Page 29; 106pp; English.

XX SQ The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells

XX SQ Sequence 12 AA;

Query Match 66.7%; Score 52; DB 2; Length 12;
 Best Local Similarity 75.0%; Pred. No. 0.56;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12

|||||

Db 1 CTLREWLSYGTC 12

RESULT 29

ID AAW25951 standard; peptide; 12 AA.

XX AC AAW25951;

XX DT 17-DEC-2001 (first entry)

XX DE Human thrombopoietin receptor (TPO-R) activator peptide #137.

XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX OS Homo sapiens.

XX PN US6251864-B1.

XX PD 26-JUN-2001.

XX PF 01-MAR-2000; 2000US-00516704.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PR 07-JUN-1996; 96WO-US009623.

XX PR 15-AUG-1996; 96US-00699027.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwirle SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;

XX DR WPI; 2001-564142/63.

XX PT Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX PS Disclosure; Col 22; 128pp; English.

XX SQ Sequences AAW25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines

XX SQ Sequence 12 AA;

Query Match 66.7%; Score 52; DB 4; Length 12;
 Best Local Similarity 75.0%; Pred. No. 0.56;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12

```
Db      1 CTLREWSYGTC 12
|||||
RESULT 30
AAW36732
ID AAW36732 standard; peptide; 13 AA.
XX
AC AAW36732;
XX
DT 09-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00484090.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
PD WPI; 1997-051883/05.
XX
PF Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
XX receptor (TR). The compound can be used for treating patients suffering
XX from haematological disorders and thrombocytopenia resulting from
XX chemotherapy, radiation therapy or bone marrow transfusions. The peptide
XX may also be used to maintain the proliferation and growth of TPO-
XX dependent cell lines and for use in biological research, for detecting
XX TPO receptors on living cells
XX
SQ Sequence 13 AA;
Query Match 66.7%; Score 52; DB 2; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CTLREWLHGGFC 12
Db 2 CTLREWSYGTC 13
|||||
RESULT 31
AAW09566
ID AAW09566 standard; protein; 10 AA.
XX
AC AAW09566;
XX
DT 10-SEP-1997 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
OS Synthetic.
XX
FN WO9640189-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US008998.
XX
PR 07-JUN-1995; 95US-00472371.
XX
PR 07-JUN-1995; 95US-00473604.
XX
PR 07-JUN-1995; 95US-00476168.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00484090.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
PD WPI; 1997-051883/05.
XX
PF Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX mimetic(s) - useful in treatment of haematological disorders, esp.
XX thrombocytopenia resulting from chemotherapy, etc.
XX
PS Disclosure; Page 29; 106pp; English.
XX
CC The present sequence is a peptide which binds to thrombopoietin (TPO)
XX receptor (TR). The compound can be used for treating patients suffering
XX from haematological disorders and thrombocytopenia resulting from
XX chemotherapy, radiation therapy or bone marrow transfusions. The peptide
XX may also be used to maintain the proliferation and growth of TPO-
XX dependent cell lines and for use in biological research, for detecting
XX TPO receptors on living cells
XX
SQ Sequence 10 AA;
Query Match 62.8%; Score 49; DB 2; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTLREWLHGG 9
Db 1 CTLREWLHGG 9
|||||
RESULT 32
AAW36717
ID AAW36717 standard; peptide; 10 AA.
XX
AC AAW36717;
XX
DT 27-FEB-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
```

PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-052226/05.
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX Disclosure; Page 29; 106pp; English.
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX Sequence 10 AA;
 SQ Query Match 62.8%; Score 49; DB 2; Length 10;
 Best Local Similarity 88.9%; Pred. No. 1.3;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTLREWLHG 9
 DB 1 CTLREWLHG 9
 RESULT 33
 AAU25936
 ID AAU25936 standard; peptide; 10 AA.
 XX AC AAU25936;
 XX 17-DEC-2001 (first entry)
 XX Human thrombopoietin receptor (TPO-R) activator peptide #122.
 DE Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX Homo sapiens.
 OS US6251864-B1.
 XX 26-JUN-2001.
 XX 01-MAR-2000; 2000US-00516704.
 XX 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US009623.
 PR 15-AUG-1996; 96US-00699027.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barret RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Poddaturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX Disclosure; Col 22; 128pp; English.
 PS Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX Sequence 10 AA;
 SQ Query Match 62.8%; Score 49; DB 4; Length 10;
 Best Local Similarity 88.9%; Pred. No. 1.3;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTLREWLHG 9
 DB 1 CTLREWLHG 9
 RESULT 34
 AAB17002
 ID AAB17002 standard; peptide; 12 AA.
 XX AC AAB17002;
 XX 31-OCT-2000 (first entry)
 XX TPO-mimetic peptide sequence SEQ ID NO:58.
 XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX Synthetic.
 OS WO200024782-A2.
 XX 04-MAY-2000.
 XX 25-OCT-1999; 99WO-US025044.
 XX 23-OCT-1998; 98US-0105371P.
 PR 22-OCT-1999; 99US-00428082.
 XX (AMGE-) AMGEN INC.
 XX Feige U, Liu C, Cheetham J, Boone TC;
 PI WPI; 2000-350702/30.
 XX Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.

PS Claim 19; Page 214; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)-a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each

CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active

CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1

CC of a and b is 1. The composition can have cytotoxic, antiaesthetic,

CC thrombolytic and immunosuppressive activities. DNAs, vectors and host

CC cells from the present invention can be used for producing pharmaceutical

CC compositions. The compositions are useful for treating cancer, asthma,

CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

CC a Fab domain) can provide a longer half-life or incorporate functions

CC such as Fc receptor binding, protein A binding, complement fixation, and

CC possibly placental transfer. AAA69443 to AAA69526 and AAA16955 to

CC AAA18003 represent nucleotide and amino acid sequences used in the

CC exemplification of the present invention

XX Sequence 12 AA;

Query Match 61.5%; Score 48; DB 3; Length 12;

Best Local Similarity 66.7%; Pred. No. 2.2;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12

Db 1 CTLREWLXXXXC 12

RESULT 35

ABB72888

ID ABB72888 standard; peptide; 12 AA.

XX ABB72888;

AC ABB72888;

XX 05-APR-2002 (first entry)

DT TPO mimetic peptide SEQ ID NO:58.

DE Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;

KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;

KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

KW cytostatic; antineumatic; antiarthritis; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antifertility; haemostatic; dermatological;

KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;

KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

KW sleep disorder; neurological degenerative disease; anaemia;

KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;

KW Fanconi's syndrome.

XX Homo sapiens.

OS Synthetic.

OS WO200193525-A2.

PN 08-NOV-2001.

PD 02-MAY-2001; 2001WO-US014310.

XX 03-MAY-2000; 2000US-00563286.

PR (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JW;

PI WPI, 2002-130313/17.

XX DR

XX Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antifertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The EPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX Sequence 12 AA;

Query Match 61.5%; Score 48; DB 5; Length 12;

Best Local Similarity 66.7%; Pred. No. 2.2;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGGFC 12

Db 1 CTLREWLXXXXC 12

RESULT 36

ADJ52674

ID ADJ52674 standard; peptide; 12 AA.

XX ADJ52674;

AC ADJ52674;

XX 06-MAY-2004 (first entry)

DT CH1 deleted mimetibody-related peptide SeqID493.

DE CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiac;

KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;

KW fungicide; gene therapy; immune disorder; cardiovascular disease;

KW arrhythmia; hypertension; heart failure; neurodegenerative;

KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;

KW cancerous condition; infectious disease; bacterial infection;

KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1..12

FT /note= "All Xaa's in this sequence are unidentified amino

FT acids"

XX WO2004002417-A2.

PN 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US020347.

XX

PR 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;
PI
XX
XX
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
XX Claim 2; SEQ ID NO 493; 129pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
XX Sequence 12 AA;
SQ
Query Match 61.5%; Score 48; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CTTLEWLHGFGC 12
Db 1 CTTLEWLXXXXC 12
RESULT 37
ADJ51635
ID ADJ51635 standard; peptide; 12 AA.
XX
XX ADJ51635;
AC
XX 06-MAY-2004 (first entry)
DT
XX CHI deleted mimetibody-related peptide SeqID493.
DE
XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
XX Unidentified.
OS
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1. .12
FT

FT
XX
XX
XX WO2004002424-A2.
XX
XX 08-JAN-2004.
XX
XX 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX 19-SEP-2002; 2002US-0412144P.
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski KA;
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 14; SEQ ID NO 493; 123pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
XX Sequence 12 AA;
SQ
Query Match 61.5%; Score 48; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CTTLEWLHGFGC 12
Db 1 CTTLEWLXXXXC 12
RESULT 38
AAW09570
ID AAW09570 standard; protein; 12 AA.
XX
XX AAW09570;
AC
XX 10-SEP-1997 (first entry)
DT
XX Thrombopoietin receptor binding peptide.
XX
XX Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX
XX Synthetic.
XX
XX WO9640189-A1.
XX
XX

XX 19-DEC-1996.
 XX 05-JUN-1996; 96WO-US008998.
 XX 07-JUN-1995; 95US-00472371.
 XX 07-JUN-1995; 95US-00473604.
 XX 07-JUN-1995; 95US-00476168.
 XX 07-JUN-1995; 95US-00478128.
 XX 07-JUN-1995; 95US-00484090.
 XX 07-JUN-1995; 95US-00485301.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-051883/05.
 XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.
 PT thrombocytopenia resulting from chemotherapy, etc.
 XX Disclosure; Page 29; 106pp; English.
 XX The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells
 XX
 XX Sequence 12 AA;
 Query Match 60.3%; Score 47; DB 2; Length 12;
 Best Local Similarity 58.3%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 CTLREWLHGGFC 12
 Db 1 CTLRQWLGDWC 12
 RESULT 39
 ID AAW36721 standard; peptide; 12 AA.
 AC AAW36721;
 DT 09-MAR-1998 (first entry)
 DE Thrombopoietin receptor binding peptide.
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 XX WO9640750-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US009623.
 XX 07-JUN-1995; 95US-00478128.
 XX 07-JUN-1995; 95US-00485301.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-052226/05.
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX Disclosure; Page 29; 106pp; English.
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 XX Sequence 12 AA;
 Query Match 60.3%; Score 47; DB 2; Length 12;
 Best Local Similarity 58.3%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 CTLREWLHGGFC 12
 Db 1 CTLRQWLGDWC 12
 RESULT 40
 ID AAU25940 standard; peptide; 12 AA.
 AC AAU25940;
 DT 17-DEC-2001 (first entry)
 DE Human thrombopoietin receptor (TPO-R) activator peptide #126.
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.
 XX Homo sapiens.
 OS US6251864-B1.
 XX 26-JUN-2001.
 XX 01-MAR-2000; 2000US-00516704.
 XX 07-JUN-1995; 95US-00478128.
 XX 07-JUN-1995; 95US-00485301.
 XX 15-AUG-1996; 96WO-US009623.
 XX 15-AUG-1996; 96US-00699027.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX Disclosure; Col 22; 128pp; English.

PI Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-052226/05.
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX Disclosure; Page 29; 106pp; English.
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 XX Sequence 12 AA;
 Query Match 60.3%; Score 47; DB 2; Length 12;
 Best Local Similarity 58.3%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 CTLREWLHGGFC 12
 Db 1 CTLRQWLGDWC 12
 RESULT 40
 ID AAU25940 standard; peptide; 12 AA.
 AC AAU25940;
 DT 17-DEC-2001 (first entry)
 DE Human thrombopoietin receptor (TPO-R) activator peptide #126.
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.
 XX Homo sapiens.
 OS US6251864-B1.
 XX 26-JUN-2001.
 XX 01-MAR-2000; 2000US-00516704.
 XX 07-JUN-1995; 95US-00478128.
 XX 07-JUN-1995; 95US-00485301.
 XX 15-AUG-1996; 96WO-US009623.
 XX 15-AUG-1996; 96US-00699027.
 XX (GLAX) GLAXO GROUP LTD.
 XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX Disclosure; Col 22; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 SQ Sequence 12 AA;

Query Match 60.3%; Score 47; DB 4; Length 12;
 Best Local Similarity 58.3%; Pred. No. 3.2;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTLRWLHGGFC 12
 ||||:|:
 Db 1 CTLRWLGDAMC 12

RESULT 41
 AAW09507
 ID AAW09507 standard; protein; 10 AA.

XX AAW09507;
 XX
 XX 10-SEP-1997 (first entry)
 XX
 DE Thrombopoietin receptor binding peptide.
 KW Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.
 XX WO9640189-A1.
 XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008998.
 XX 07-JUN-1995; 95US-00472371.
 PR 07-JUN-1995; 95US-00473604.
 PR 07-JUN-1995; 95US-00476168.
 PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00484090.
 PR 07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.
 PT thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 27; 106pp; English.

XX The present sequence is a peptide which binds to thrombopoietin (TPO)

CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells
 XX
 SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 2; Length 10;
 Best Local Similarity 77.8%; Pred. No. 3.8;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTLRWLHGG 9
 ||||:|:
 Db 1 CTLRWLQGG 9

RESULT 42
 AAW36658
 ID AAW36658 standard; peptide; 10 AA.

XX AAW36658;
 AC AAW36658;
 XX 11-MAR-1998 (first entry)

DT Thrombopoietin receptor binding peptide.
 DE
 XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX Synthetic.
 XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.
 XX 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX WPI; 1997-052226/05.
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 27; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 2; Length 10;
 Best Local Similarity 77.8%; Pred. No. 3.8;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTLRWLHGG 9

```

Db      1 CTLRWLQG 9
||||:|
RESULT 43
AAB16990
ID AAB16990 standard; peptide; 10 AA.
XX
AC AAB16990;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:46.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 211; 609pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
XX Sequence 10 AA;
Query Match 59.0%; Score 46; DB 3; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTLRWLHG 9
|||:|
Db 1 CTLRWLQG 9

```

```

Db      1 CTLRWLQG 9
||||:|
RESULT 44
AAU25877
ID AAU25877 standard; peptide; 10 AA.
XX
AC AAU25877;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #63.
XX
KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
PF 01-MAR-2000; 2000US-00516704.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ,
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
DR WPI; 2001-564142/63.
XX
PT Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
PS Disclosure; Col 20; 128pp; English.
XX
CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 10 AA;
Query Match 59.0%; Score 46; DB 4; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTLRWLHG 9
|||:|
Db 1 CTLRWLQG 9

```


RESULT 45
ABB72876
ID ABB72876 standard; peptide; 10 AA.
XX
AC ABB72876;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:46.
XX
KW Modified peptide: mimetic; Fc domain; fusion; immunoglobulin G; IGG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX WPI; 2002-130313/17.
XX
DR Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 5; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTLREWLHG 9
DB 1 CTLRWLQGG 9

RESULT 46
ADJ73027
ID ADJ73027 standard; peptide; 10 AA.
XX
AC ADJ73027;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 481.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallion BJ, Ghraieb J;
XX WPI; 2003-804237/75.
XX
DR New CDR mimetibody comprising a portion of a heavy or light chain
XX variable region comprising human framework or ligand binding region,
XX useful for preparing a composition for treating e.g., immune,
XX cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 481; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 7; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTLREWLHG 9
DB 1 CTLRWLQGG 9

RESULT 47

ADJ52662
 ID ADJ52662 standard; peptide; 10 AA.
 XX
 AC ADJ52662;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE CH1 deleted mimetibody-related peptide SeqID481.
 XX
 KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiac;
 KW hypotensive; neuroprotective; notropic; antibacterial; viricide;
 KW fungicide; gene therapy; immune disorder; cardiovascular disease;
 KW arrhythmia; hypertension; heart failure; neurodegenerative;
 KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 KW cancerous condition; infectious disease; bacterial infection;
 KW viral infection; fungal infection.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 FN WO2004002417-A2.
 XX
 PD 08-JAN-2004.
 XX
 PF 27-JUN-2003; 2003WO-US020347.
 XX
 PR 28-JUN-2002; 2002US-0392431P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutolowski KA;
 XX
 DR WPI, 2004-082870/08.
 XX
 XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
 PT modulating, treating, alleviating, preventing an immune, cardiovascular,
 PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
 PT diseases.
 XX
 PS Claim 2; SEQ ID NO 481; 129pp; English.
 XX
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, notropic,
 CC antibacterial, viricide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.
 XX
 SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 8; Length 10;
 Best Local Similarity 77.8%; Pred. NO. 3.8;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CTLREWHLG 9
 |||||
 Db 1 CTLRWLQGG 9

RESULT 48

ADJ51623
 ID ADJ51623 standard; peptide; 10 AA.
 XX
 AC ADJ51623;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE CH1 deleted mimetibody-related peptide SeqID481.
 XX
 KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen; dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen; gynaecological-Gen; hepatotropic; haemostatic; immunomodulator; antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic; ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor; TNF; cytokines; bone disorder; joint disorder; cardiovascular disorder; dental disorder; oral disorder; dermatological disorder; ear disorder; nose disorder; throat disorder; endocrine disorder; metabolic disorder; gastrointestinal disorder; gynaecological disorder; hepatic disorder; obstetric disorder; haematologic disorder; immunological disorder; allergic disorder; infectious disorder; musculoskeletal disorder; oncological disorder; neurological disorder; nutritional disorder; ophthalmologic disorder; pediatric disorder; psychiatric disorder; renal disorder; pulmonary disorder.
 XX
 OS Unidentified.
 OS Synthetic.
 XX
 FN WO2004002424-A2.
 XX
 PD 08-JAN-2004.
 XX
 PF 30-JUN-2003; 2003WO-US020495.
 XX
 PR 28-JUN-2002; 2002US-0392431P.
 PR 19-SEP-2002; 2002US-0412144P.
 XX
 PA (CENZ) CENTOCOR INC.
 XX
 PI Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 PI Kutolowski KA;
 XX
 DR WPI, 2004-082872/08.
 XX
 XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
 PT nutritional disorders.
 XX
 PS Claim 14; SEQ ID NO 481; 123pp; English.
 XX
 CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen, gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic, immunomodulator, antiallergic, muscular-Gen, cytostatic, antiinflammatory, neuroleptic, ophthalmological, nephrotropic or respiratory-Gen activity acting as a tumour necrosis factor (TNF)-modulator or cytokine-agonist. The methods and compositions of the present invention are useful for the diagnosis, prevention and/or treatment of diseases or conditions associated with aberrant expression or activity of the CHI deleted mimetibody, such as a bone or joint, cardiovascular, dental or oral, dermatological, ear, nose or throat, endocrine, metabolic, gastrointestinal, gynaecological, hepatic, obstetric, haematologic, immunological, allergic, infectious, musculoskeletal, oncological, neurological, nutritional, ophthalmologic, pediatric, psychiatric, renal or pulmonary disorders. The present sequence is that of a peptide which may be used during the creation of a mimetibody of the invention.
 XX
 SQ Sequence 10 AA;

Query Match 59.0%; Score 46; DB 8; Length 10;

```

Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHG 9
DB 1 CTLRWLQGS 9

RESULT 49
AAW09523
ID AAW09523 standard; protein; 12 AA.
XX AC
XX AC AAW09523;
XX DT 10-SEP-1997 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Haematology; thrombocytopenia; TPO; TR; proliferation;
XX KW bone marrow transfusion; chemotherapy; radiation therapy.
XX OS Synthetic.
XX WO9640189-A1.
XX PD 19-DEC-1996.
XX PF 05-JUN-1996; 96WO-US008998.
XX PR 07-JUN-1995; 95US-00472371.
XX PR 07-JUN-1995; 95US-00473604.
XX PR 07-JUN-1995; 95US-00476168.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00484090.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
XX PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX DR WPI; 1997-051883/05.
XX DT Thrombopoietin receptor-binding/activating peptide(s) and peptide
XX PT mimetic(s) - useful in treatment of haematological disorders, esp.
XX PT thrombocytopenia resulting from chemotherapy, etc.
XX PS Disclosure; Page 27; 106pp; English.
XX CC The present sequence is a peptide which binds to thrombopoietin (TPO)
XX CC receptor (TR). The compound can be used for treating patients suffering
XX CC from haematological disorders and thrombocytopenia resulting from
XX CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
XX CC may also be used to maintain the proliferation and growth of TPO-
XX CC dependent cell lines and for use in biological research, for detecting
XX CC TPO receptors on living cells
XX SQ Sequence 12 AA;
Query Match 59.0%; Score 46; DB 2; Length 12;
Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
DB 1 CTLRWILLGMC 12

Search completed: May 12, 2006, 10:35:24
Job time : 118.279 secs

27-FEB-1998 (first entry)
Thrombopoietin receptor binding peptide.
Thrombopoietin receptor; binding peptide; treatment; agonist;
haematological disorder; thrombocytopenia; chemotherapy;
radiation therapy; bone marrow transfusion; diagnosis;
signal transduction; receptor activation; cell culture.
Synthetic.
WO9640750-A1.
19-DEC-1996.
07-JUN-1996; 96WO-US009623.
07-JUN-1995; 95US-00478128.
07-JUN-1995; 95US-00485301.
(GLAX ) GLAXO GROUP LTD.
Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
WPI; 1997-052226/05.
Peptides and peptide mimetics which bind to and activate the
thrombopoietin receptor - useful in treatment of haematological
disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
Disclosure; Page 27; 106pp; English.
The present peptide, which binds the thrombopoietin receptor (TR), can be
used to treat disorders which are susceptible to treatment with a
thrombopoietin agonist, preferably haematological disorders and
thrombocytopenia resulting from chemotherapy, radiation therapy or bone
marrow transfusions. It can also be used diagnostically, e.g. to
investigate the mechanism of thrombopoietin signal transduction and
receptor activation, or to maintain the proliferation and growth of
thrombopoietin dependent cell lines
XX SQ Sequence 12 AA;
Query Match 59.0%; Score 46; DB 2; Length 12;
Best Local Similarity 58.3%; Pred. No. 4.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
DB 1 CTLRWILLGMC 12

Search completed: May 12, 2006, 10:35:24
Job time : 118.279 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 14.5641 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTLREWLHGGFC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

PIR 80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	60.3	97	1	cytochrome-c oxida
2	45	57.7	361	2	hypothetical prote
3	45	57.7	503	2	hypothetical prote
4	45	57.7	532	1	carboxypeptidase C
5	44	56.4	171	2	hypothetical prote
6	43	55.1	791	2	diacylglycerol kin
7	42	53.8	380	2	ubiquinol-cytochro
8	41	52.6	218	2	conserved hypothet
9	41	52.6	218	2	CT383 hypothetical
10	41	52.6	309	2	FeD8.20 [imported]
11	41	52.6	530	2	probable peramease
12	41	52.6	531	2	ABC transporter, p
13	41	52.6	619	2	hypothetical prote
14	41	52.6	1025	2	hypothetical prote
15	41	52.6	2225	1	pyrimidine synthe
16	41	52.6	4342	2	probable non-ribos
17	40.5	51.9	4391	2	parlecan precursor
18	40	51.3	117	2	hypothetical prote
19	40	51.3	139	2	hypothetical prote
20	40	51.3	312	2	hypothetical prote
21	40	51.3	378	2	ubiquinol-cytochro
22	40	51.3	379	1	ubiquinol-cytochro
23	40	51.3	379	1	ubiquinol-cytochro
24	40	51.3	379	1	ubiquinol-cytochro
25	40	51.3	379	1	ubiquinol-cytochro
26	40	51.3	379	1	ubiquinol-cytochro
27	40	51.3	379	1	ubiquinol-cytochro
28	40	51.3	379	1	ubiquinol-cytochro
29	40	51.3	379	1	ubiquinol-cytochro

30	40	51.3	379	1	S43262	ubiquinol-cytochro
31	40	51.3	379	1	S43263	ubiquinol-cytochro
32	40	51.3	379	1	S43264	ubiquinol-cytochro
33	40	51.3	379	1	S43265	ubiquinol-cytochro
34	40	51.3	379	1	S43266	ubiquinol-cytochro
35	40	51.3	379	1	S43267	ubiquinol-cytochro
36	40	51.3	379	1	S43269	ubiquinol-cytochro
37	40	51.3	379	1	S43269	ubiquinol-cytochro
38	40	51.3	379	2	E58889	ubiquinol-cytochro
39	40	51.3	379	2	I48132	ubiquinol-cytochro
40	40	51.3	379	2	I48132	ubiquinol-cytochro
41	40	51.3	379	2	I48180	ubiquinol-cytochro
42	40	51.3	379	2	E58851	ubiquinol-cytochro
43	40	51.3	379	2	T11375	ubiquinol-cytochro
44	40	51.3	379	2	T11869	ubiquinol-cytochro
45	40	51.3	379	2	T11492	ubiquinol-cytochro
46	40	51.3	379	2	S58990	ubiquinol-cytochro
47	40	51.3	379	2	T11401	ubiquinol-cytochro
48	40	51.3	379	2	A53077	ubiquinol-cytochro
49	40	51.3	379	2	I48134	ubiquinol-cytochro
50	40	51.3	380	1	CBRT	ubiquinol-cytochro
51	40	51.3	381	1	CBMS	ubiquinol-cytochro
52	40	51.3	382	1	T11440	ubiquinol-cytochro
53	40	51.3	382	1	S33573	ubiquinol-cytochro
54	40	51.3	382	2	S47882	ubiquinol-cytochro
55	40	51.3	438	2	C86244	DnaJ homolog, 4706
56	40	51.3	2180	2	T29764	hypothetical prote
57	40	51.3	2321	2	S45306	notch 3 protein -
58	40	51.3	2321	2	S78549	notch3 protein - h
59	39	50.0	4568	2	T08030	dynein beta heavy
60	39	50.0	160	2	H83047	hypothetical prote
61	39	50.0	209	2	A13455	transcription regu
62	39	50.0	217	2	I48902	homeobox protein P
63	39	50.0	252	2	T29527	hypothetical prote
64	39	50.0	269	2	F71519	probable rRNA meth
65	39	50.0	372	2	T11100	ubiquinol-cytochro
66	39	50.0	379	1	S43268	ubiquinol-cytochro
67	39	50.0	379	1	S43270	ubiquinol-cytochro
68	39	50.0	384	2	S46110	hypothetical prote
69	39	50.0	492	2	T01086	probable serine/th
70	39	50.0	545	2	B44054	orf2 protein - Jun
71	39	50.0	726	2	T44187	infected cell prot
72	39	50.0	726	2	T44000	transport protein
73	39	50.0	861	2	A48825	Notch homolog Motc
74	39	50.0	899	2	G02428	subtilisin-like pr
75	39	50.0	915	1	A48225	subtilisin-like pr
76	39	50.0	915	1	A48225	subtilisin-like pr
77	39	50.0	915	2	JC6148	probable propotei
78	39	50.0	1472	2	B48225	alpha-2-macroglobu
79	39	50.0	1474	1	MAHU	alpha-2-macroglobu
80	39	50.0	1476	2	JC5143	pregnancy zone pro
81	39	50.0	1482	2	S13495	notch protein homo
82	39	50.0	2531	2	S18188	notch-1 protein -
83	39	50.0	2531	2	A46019	methionine-trna sy
84	38.5	49.4	689	2	C84192	hypothetical prote
85	38	48.7	112	2	D95269	T-cell receptor al
86	38	48.7	141	2	D32536	hypothetical prote
87	38	48.7	204	2	C83748	hypothetical prote
88	38	48.7	227	2	T49725	hypothetical prote
89	38	48.7	285	2	T22722	conserved hypothet
90	38	48.7	312	2	A87449	hypothetical prote
91	38	48.7	338	2	JC7161	acyl-CoA hydrolase
92	38	48.7	338	2	JC5415	palmitoyl-CoA hydr
93	38	48.7	343	2	JC5416	palmitoyl-CoA hydr
94	38	48.7	442	2	H86144	hypothetical prote
95	38	48.7	445	2	E84714	probable protein k
96	38	48.7	482	2	G86227	hypothetical prote
97	38	48.7	497	2	T15812	hypothetical prote
98	38	48.7	506	2	H75522	probable carboxyle
99	38	48.7	555	2	S63177	mannosyl transfera
100	38	48.7	631	2	T05103	hypothetical prote
101	38	48.7	658	1	A65079	arginine decarboxy
102	38	48.7	658	2	AG0876	arginine decarboxy

103	38	48.7	658	2	F91105	176	37	47.4	1023	2	E71376	conserved hypothet
104	38	48.7	658	2	A85951	177	37	47.4	1370	2	T19188	hypothetical prote
105	38	48.7	659	2	A80114	178	37	47.4	1378	2	G88637	protein F5H1.4 li
106	38	48.7	695	2	E87397	179	37	47.4	3229	2	S27852	probable cell-surf
107	38	48.7	717	2	T28829	180	36.5	46.8	203	2	T24042	hypothetical prote
108	38	48.7	749	2	T31536	181	36.5	46.8	452	2	AG1233	cohyrnic acid a,c
109	38	48.7	1057	2	T25396	182	36	46.2	100	2	S76498	ribosomal protein
110	38	48.7	1444	2	B84809	183	36	46.2	100	2	AB2302	30S ribosomal prot
111	37.5	48.1	1274	2	E81779	184	36	46.2	120	2	B90158	conserved hypothet
112	37	47.4	85	2	S01157	185	36	46.2	178	2	T29352	hypothetical prote
113	37	47.4	97	2	I49360	186	36	46.2	190	2	PQ0019	adenosinetriphosph
114	37	47.4	98	2	AC2369	187	36	46.2	220	2	D87535	glutathione S-tran
115	37	47.4	140	2	E83910	188	36	46.2	225	2	B72749	probable proteasom
116	37	47.4	167	2	I50487	189	36	46.2	233	2	S52461	hypothetical prote
117	37	47.4	170	2	B97811	190	36	46.2	233	2	S38228	hypothetical prote
118	37	47.4	214	2	B1247	191	36	46.2	233	2	S53104	hypothetical prote
119	37	47.4	218	2	G82019	192	36	46.2	279	2	S75971	hypothetical prote
120	37	47.4	230	2	AB3012	193	36	46.2	279	2	S62364	L-2-chloropropioni
121	37	47.4	230	2	E98272	194	36	46.2	301	2	S62416	n-acetylglucosamin
122	37	47.4	297	1	S49348	195	36	46.2	333	2	E97257	spore coat protein
123	37	47.4	302	2	JH0572	196	36	46.2	345	2	T29786	hypothetical prote
124	37	47.4	308	1	S22931	197	36	46.2	345	2	T16869	hypothetical prote
125	37	47.4	308	1	S22931	198	36	46.2	370	2	J80342	E2F transcription
126	37	47.4	308	2	S22921	199	36	46.2	377	2	C70628	hypothetical prote
127	37	47.4	308	2	S22925	200	36	46.2	379	2	I49359	ubiquinol-cytochro
128	37	47.4	308	2	S22929	201	36	46.2	427	2	S74211	PAS-6/7 protein pr
129	37	47.4	308	2	S22924	202	36	46.2	430	2	B83633	hypothetical prote
130	37	47.4	308	2	S22927	203	36	46.2	454	2	A10725	para-aminobenzoate
131	37	47.4	308	2	S22930	204	36	46.2	454	2	A31132	p-aminobenzoate sy
132	37	47.4	308	2	S22920	205	36	46.2	470	2	E83947	hypothetical prote
133	37	47.4	308	2	S22928	206	36	46.2	471	2	D83546	probable amino aci
134	37	47.4	308	2	S22922	207	36	46.2	472	2	G83537	aromatic amino aci
135	37	47.4	320	2	I51030	208	36	46.2	473	2	B81007	cysteinyI-tRNA syn
136	37	47.4	332	2	A39711	209	36	46.2	520	2	T37957	probable cytochrom
137	37	47.4	339	2	C86978	210	36	46.2	617	2	S08317	hydrogenase (BC 1.
138	37	47.4	339	2	F70901	211	36	46.2	619	2	A45625	phosphoenolpyruvat
139	37	47.4	344	2	A48990	212	36	46.2	624	2	A83237	hypothetical prote
140	37	47.4	379	2	D90617	213	36	46.2	628	2	A55421	nucleoside-triphos
141	37	47.4	379	2	D90621	214	36	46.2	629	2	T37255	acetylcholinestera
142	37	47.4	379	2	T10998	215	36	46.2	646	2	S36586	E1 protein - human
143	37	47.4	379	2	D90625	216	36	46.2	666	2	T17396	vrlp protein - Dic
144	37	47.4	379	2	D90615	217	36	46.2	672	2	JC4637	transketolase (BC
145	37	47.4	379	2	T11178	218	36	46.2	672	2	H03457	transketolase (BC
146	37	47.4	379	2	T11530	219	36	46.2	699	2	H82030	cysteine-tRNA liga
147	37	47.4	379	2	D90613	220	36	46.2	788	2	T25061	hypothetical prote
148	37	47.4	379	2	D90627	221	36	46.2	788	2	T29768	hypothetical prote
149	37	47.4	379	2	D90619	222	36	46.2	860	2	C86203	hypothetical prote
150	37	47.4	380	1	S34285	223	36	46.2	973	2	AB2340	hypothetical prote
151	37	47.4	380	1	S04840	224	36	46.2	1119	2	T37460	probable sodium bi
152	37	47.4	380	1	S10198	225	36	46.2	1175	2	T22491	hypothetical prote
153	37	47.4	380	2	T11033	226	36	46.2	1201	2	F81202	proline dehydrogen
154	37	47.4	380	2	T11803	227	36	46.2	1230	2	T07663	soluble starch syn
155	37	47.4	380	2	I51374	228	36	46.2	1566	2	T20058	hypothetical prote
156	37	47.4	380	2	E58893	229	36	46.2	1584	2	T18276	protein-tyrosine k
157	37	47.4	380	2	T11466	230	36	46.2	1674	2	T01265	starch synthase DU
158	37	47.4	380	2	T11204	231	36	46.2	2531	2	H31070	notch homolog - se
159	37	47.4	380	2	T11191	232	35.5	45.5	292	2	H82955	hypothetical prote
160	37	47.4	380	2	S42245	233	35.5	45.5	298	2	G72531	hypothetical prote
161	37	47.4	380	2	T11086	234	35.5	45.5	394	2	A48100	B-type cyclin, Cig
162	37	47.4	381	2	S68140	235	35.5	45.5	411	2	S44344	cyclin cycl17 - fis
163	37	47.4	384	2	S34816	236	35.5	45.5	753	2	T19338	hypothetical prote
164	37	47.4	420	2	T36193	237	35.5	45.5	832	2	E84543	probable beta-gala
165	37	47.4	437	2	T18555	238	35	44.9	101	2	F72573	hypothetical prote
166	37	47.4	477	2	A75052	239	35	44.9	120	2	S10865	early E4 13K prote
167	37	47.4	591	2	G01586	240	35	44.9	145	2	S74292	hypothetical prote
168	37	47.4	594	2	E96667	241	35	44.9	169	2	T31484	hypothetical prote
169	37	47.4	644	2	A43370	242	35	44.9	174	2	G83712	hypothetical prote
170	37	47.4	701	2	S61239	243	35	44.9	185	2	F35719	hsp-like nucleoti
171	37	47.4	754	2	S04381	244	35	44.9	185	2	B86104	ATP-binding compon
172	37	47.4	850	2	S56015	245	35	44.9	185	2	E91263	tumor-associated a
173	37	47.4	981	2	T50851	246	35	44.9	197	2	JC6544	ribosomal protein
174	37	47.4	984	2	E70406	247	35	44.9	206	2	D97285	hypothetical prote
175	37	47.4	987	2	T50850	248	35	44.9	216	2	T18176	

249	35	44.9	228	2	JC4869	ribonuclease S4 (E	322	35	44.9	429	2	AD3642	(S)-2-hydroxy-acid
250	35	44.9	229	2	F70897	hypothetical prote	323	35	44.9	433	2	AE2658	glycolate oxidase
251	35	44.9	230	2	T46203	transcription fact	324	35	44.9	433	2	B97440	glycolate oxidase
252	35	44.9	242	2	G91139	probable peptidogl	325	35	44.9	442	2	T01731	hypothetical prote
253	35	44.9	242	2	B85985	probable peptidogl	326	35	44.9	449	2	C69079	probable phosphoma
254	35	44.9	242	2	B65112	Monofunctional bio	327	35	44.9	462	2	B81352	cysteine-tRNA liga
255	35	44.9	243	2	S58236	hypothetical prote	328	35	44.9	462	2	H69751	amino acid permeas
256	35	44.9	243	2	C95908	hypothetical prote	329	35	44.9	466	2	G83663	cysteinyl-tRNA syn
257	35	44.9	275	2	T10310	apoptosis-inhibiti	330	35	44.9	466	2	AG0839	GABA permease (4-a
258	35	44.9	281	2	A43749	lignin beta-ether	331	35	44.9	466	2	A39599	55k erythrocyte me
259	35	44.9	299	2	B82535	conserved hypothet	332	35	44.9	467	2	F83009	probable amino aci
260	35	44.9	305	2	A46476	B cell-associated	333	35	44.9	481	2	T05270	probable serine/th
261	35	44.9	306	1	CPBOB	carboxypeptidase B	334	35	44.9	494	2	D84860	probable protein k
262	35	44.9	308	2	S22932	ubiquinol-cytochro	335	35	44.9	495	2	B96609	probable protein k
263	35	44.9	308	2	S22932	ubiquinol-cytochro	336	35	44.9	502	2	AE0666	probable esterase
264	35	44.9	312	2	H88567	protein K11H3.3 [i	337	35	44.9	502	2	T35910	probable carboxyle
265	35	44.9	315	2	B98226	hypothetical prote	338	35	44.9	502	2	A39351	tetragonal surface
266	35	44.9	315	2	AF3060	conserved hypothet	339	35	44.9	512	2	T47793	receptor-like prot
267	35	44.9	341	2	T07145	epoxide hydrolase	340	35	44.9	515	1	VHNZP3	hypothetical prote
268	35	44.9	342	2	AD3450	proline racemase (341	35	44.9	517	2	T18226	nucleocapsid prote
269	35	44.9	351	2	I39808	spore coat protein	342	35	44.9	524	1	A48341	nucleocapsid prote
270	35	44.9	361	2	T45938	hypothetical prote	343	35	44.9	563	2	S54420	invasion protein 1
271	35	44.9	366	2	T11271	ubiquinol-cytochro	344	35	44.9	604	2	G89864	hypothetical prote
272	35	44.9	368	2	T26338	hypothetical prote	345	35	44.9	608	2	F83397	probable peptidase
273	35	44.9	370	2	H64480	hypothetical prote	346	35	44.9	618	2	S09251	hydrogenase (EC 1.
274	35	44.9	374	2	S40756	ubiquinol-cytochro	347	35	44.9	623	2	B81037	lipopolysaccharide
275	35	44.9	378	1	S17412	ubiquinol-cytochro	348	35	44.9	630	2	A81196	biosynthetic argin
276	35	44.9	378	2	D83381	hypothetical prote	349	35	44.9	630	2	D81831	probable decarboxy
277	35	44.9	379	1	CBBO	ubiquinol-cytochro	350	35	44.9	673	2	T36717	probable serine/th
278	35	44.9	379	1	S17405	ubiquinol-cytochro	351	35	44.9	692	2	B64381	hypothetical prote
279	35	44.9	379	1	S17407	ubiquinol-cytochro	352	35	44.9	697	2	H85205	potassium channel
280	35	44.9	379	1	S17409	ubiquinol-cytochro	353	35	44.9	724	1	QOBEI5	U89 protein - hum
281	35	44.9	379	1	S17411	ubiquinol-cytochro	354	35	44.9	801	2	A47744	diacylglycerol kin
282	35	44.9	379	1	S17413	ubiquinol-cytochro	355	35	44.9	823	2	S59141	mdrA protein - ma
283	35	44.9	379	1	S17414	ubiquinol-cytochro	356	35	44.9	952	2	B86147	hypothetical prote
284	35	44.9	379	1	S17418	ubiquinol-cytochro	357	35	44.9	1204	2	A96676	hypothetical prote
285	35	44.9	379	1	S17419	ubiquinol-cytochro	358	35	44.9	1267	1	MXR32	lambda 3 protein -
286	35	44.9	379	1	S17420	ubiquinol-cytochro	359	35	44.9	1275	2	T38397	probable GTPase ac
287	35	44.9	379	2	I48135	ubiquinol-cytochro	360	35	44.9	1281	2	T15762	hypothetical prote
288	35	44.9	379	2	I49400	ubiquinol-cytochro	361	35	44.9	1328	2	T43060	agrin - electric r
289	35	44.9	379	2	T11453	ubiquinol-cytochro	362	35	44.9	1341	2	H98323	hypothetical prote
290	35	44.9	379	2	T11505	ubiquinol-cytochro	363	35	44.9	1426	2	E90456	oxydoreductase, pr
291	35	44.9	379	2	S58455	ubiquinol-cytochro	364	35	44.9	1444	2	T18856	angiogenesis inhib
292	35	44.9	379	2	S58454	ubiquinol-cytochro	365	35	44.9	1458	1	A49707	phospholipase A2 r
293	35	44.9	379	2	T11414	ubiquinol-cytochro	366	35	44.9	1597	1	BVEFSL	sol protein, large
294	35	44.9	379	2	S58451	ubiquinol-cytochro	367	35	44.9	1597	2	T08428	gene small optic l
295	35	44.9	379	2	S58452	ubiquinol-cytochro	368	35	44.9	1873	2	A55645	calcium channel, v
296	35	44.9	379	2	S58457	ubiquinol-cytochro	369	35	44.9	1959	1	AGRT	agrin - rat
297	35	44.9	379	2	T11259	ubiquinol-cytochro	370	35	44.9	3010	1	GNWVCJ	genome polyprotein
298	35	44.9	379	2	T11349	ubiquinol-cytochro	371	35	44.9	3010	1	S18030	genome polyprotein
299	35	44.9	379	2	S50334	ubiquinol-cytochro	372	35	44.9	3011	1	GNWVC3	genome polyprotein
300	35	44.9	379	2	S58462	ubiquinol-cytochro	373	35	44.9	3011	1	GNWVCJ	genome polyprotein
301	35	44.9	379	2	S58057	ubiquinol-cytochro	374	35	44.9	3011	1	S40770	genome polyprotein
302	35	44.9	379	2	S58085	ubiquinol-cytochro	375	35	44.9	3011	1	IUFFTM	cadherin-related t
303	35	44.9	379	2	S58456	ubiquinol-cytochro	376	35	44.9	6831	2	A88852	protein unc-22 [im
304	35	44.9	379	2	S58459	ubiquinol-cytochro	377	35	44.9	6839	2	S57242	witchin [similar
305	35	44.9	379	2	S58458	ubiquinol-cytochro	378	35	44.9	7160	2	T27935	hypothetical prote
306	35	44.9	379	2	S58458	ubiquinol-cytochro	379	35	44.9	212	2	C87585	hypothetical prote
307	35	44.9	380	2	T11299	ubiquinol-cytochro	380	34.5	44.2	212	2	C69530	3-ketoacyl-CoA thi
308	35	44.9	386	2	S72168	dopamine receptor	381	34.5	44.2	414	2	T48079	hypothetical prote
309	35	44.9	391	2	E83151	hypothetical prote	382	34.5	44.2	1401	2	T48079	hypothetical prote
310	35	44.9	395	2	F71118	conserved hypothet	383	34	43.6	91	2	G83144	hypothetical prote
311	35	44.9	397	2	AE0148	probable periplasm	384	34	43.6	97	1	OGHU6A	hypothetical prote
312	35	44.9	397	2	B81839	probable transpos	385	34	43.6	102	2	C81682	conserved hypothet
313	35	44.9	409	2	D75629	carboxypeptidase B	386	34	43.6	105	2	T35995	probable dioxygena
314	35	44.9	415	2	A32129	carboxypeptidase B	387	34	43.6	106	2	JQ0234	hypothetical 12.5K
315	35	44.9	416	1	A42332	carboxypeptidase B	388	34	43.6	114	2	T11095	NAHD2 dehydrogena
316	35	44.9	416	2	D75407	probable transpos	389	34	43.6	116	2	A71565	hypothetical prote
317	35	44.9	417	1	A34487	carboxypeptidase A	390	34	43.6	137	2	B96603	transcription fact
318	35	44.9	423	2	A41204	carboxypeptidase B	391	34	43.6	138	2	G69328	Brute force ORF -
319	35	44.9	423	2	D90927	hypothetical prote	392	34	43.6	143	2	H70601	hypothetical prote
320	35	44.9	423	2	A64926	ynhC protein - Esc	393	34	43.6	166	2	H64416	VP829-like phospho
321	35	44.9	423	2	H85775	hypothetical prote	394	34	43.6	169	2	D72649	hypothetical prote

395	34	43.6	185	2	E70029	conserved hypothet	468	34	43.6	511	2	E64183	xylokinase homol
396	34	43.6	186	2	F90190	conserved hypotchet	469	34	43.6	511	2	A69369	glutamate synthase
397	34	43.6	211	2	C82381	pyridoxamine 5'-ph	470	34	43.6	513	2	C58888	cytochrome-c oxida
398	34	43.6	243	2	G98156	hypothetical prote	471	34	43.6	513	2	T11443	cytochrome-c oxida
399	34	43.6	248	1	JQ1682	infected cell prot	472	34	43.6	513	2	T11482	cytochrome-c oxida
400	34	43.6	259	2	AG2462	hypothetical prote	473	34	43.6	513	2	T11391	cytochrome-c oxida
401	34	43.6	261	2	G88469	protein C28H8.10 [474	34	43.6	514	1	ODMG1	cytochrome-c oxida
402	34	43.6	267	2	I72882r	Fc gamma receptor	475	34	43.6	514	2	T11495	cytochrome-c oxida
403	34	43.6	267	2	I56110	Fc-gamma RIIB-alp	476	34	43.6	514	2	T11142	cytochrome-c oxida
404	34	43.6	267	2	A35902	Fc gamma (IgG) rec	477	34	43.6	514	2	T11404	cytochrome-c oxida
405	34	43.6	290	2	G65031	hypothetical prote	478	34	43.6	514	2	S41837	cytochrome-c oxida
406	34	43.6	290	2	B85899	hypothetical prote	479	34	43.6	514	2	S26153	cytochrome-c oxida
407	34	43.6	290	2	B91055	hypothetical prote	480	34	43.6	514	2	T11365	cytochrome-c oxida
408	34	43.6	292	2	T02349	hypothetical prote	481	34	43.6	514	2	T11859	cytochrome-c oxida
409	34	43.6	297	2	S55063	conserved hypothet	482	34	43.6	514	2	T11249	cytochrome-c oxida
410	34	43.6	298	2	AD2715	ABC transporter, m	483	34	43.6	514	2	T10974	cytochrome-c oxida
411	34	43.6	299	2	AH0245	probable 4-diphosp	484	34	43.6	514	2	S04749	cytochrome-c oxida
412	34	43.6	301	2	H70505	hypothetical prote	485	34	43.6	515	1	VHN2B3	nucleocapsid prote
413	34	43.6	306	2	A72668	probable spermidin	486	34	43.6	515	2	C58892	cytochrome-c oxida
414	34	43.6	308	2	S22926	ubiquinol-cytochro	487	34	43.6	515	2	T11339	cytochrome-c oxida
415	34	43.6	308	2	A84566	Ac-like transposas	488	34	43.6	516	2	S36008	cytochrome-c oxida
416	34	43.6	308	2	S58504	reverse transcript	489	34	43.6	516	2	S35464	cytochrome-c oxida
417	34	43.6	309	2	A38395	mast cell carboxyp	490	34	43.6	516	2	T09859	cytochrome-c oxida
418	34	43.6	313	2	T35935	transcription regu	491	34	43.6	516	2	T09949	cytochrome-c oxida
419	34	43.6	315	2	T18696	hypothetical prote	492	34	43.6	516	2	S45351	cytochrome-c oxida
420	34	43.6	320	2	A1892	[NiFe] uptake hydr	493	34	43.6	517	2	S55006	cytochrome-c oxida
421	34	43.6	321	2	S58686	drpGglucose 4,6-de	494	34	43.6	517	2	T11302	cytochrome-c oxida
422	34	43.6	322	1	HLHUR2	T-cell surface gly	495	34	43.6	518	2	T11536	cytochrome-c oxida
423	34	43.6	328	1	DWSMGG	drpGglucose 4,6-de	496	34	43.6	518	2	T11456	cytochrome-c oxida
424	34	43.6	334	2	I59348	CCAAT binding tran	497	34	43.6	518	2	T11766	cytochrome-c oxida
425	34	43.6	335	2	T51106	drpGglucose 4,6-de	498	34	43.6	518	2	T11289	cytochrome-c oxida
426	34	43.6	353	1	C69979	conserved hypothet	499	34	43.6	521	2	T11166	CDPdiacylglycerol-
427	34	43.6	358	1	WBBE38	infected cell prot	500	34	43.6	522	2	T45861	hypothetical prote
428	34	43.6	360	2	S34256	photosystem II pro	501	34	43.6	523	2	T22728	nucleocapsid prote
429	34	43.6	371	2	P80016	5-methyltetrahydro	502	34	43.6	524	1	VHN2P1	nucleocapsid prote
430	34	43.6	375	1	YMYI	citrate (5i)-synth	503	34	43.6	524	1	VHN2T1	hypothetical prote
431	34	43.6	378	1	C30020	ubiquinol-cytochro	504	34	43.6	544	2	B83347	hypothetical prote
432	34	43.6	378	1	S01190	ubiquinol-cytochro	505	34	43.6	547	2	D84942	methionine-tRNA li
433	34	43.6	379	2	S58466	ubiquinol-cytochro	506	34	43.6	568	1	RN235	polymerase-associa
434	34	43.6	380	2	A71390	ubiquinol-cytochro	507	34	43.6	568	1	RN239	polymerase-associa
435	34	43.6	388	1	CBZM	ubiquinol-cytochro	508	34	43.6	571	1	JN0858	chitinase (SC 3.2.
436	34	43.6	388	2	T29364	hypothetical prote	509	34	43.6	571	2	T42071	probable chitinase
437	34	43.6	392	1	CBVF	ubiquinol-cytochro	510	34	43.6	584	2	I50419	s-glycerin precurs
438	34	43.6	393	1	CBPM	ubiquinol-cytochro	511	34	43.6	597	2	AH2351	serine/threonine k
439	34	43.6	393	1	S38960	ubiquinol-cytochro	512	34	43.6	603	2	H90061	hypothetical prote
440	34	43.6	394	1	CSOBE	ubiquinol-cytochro	513	34	43.6	609	2	AH0388	alpha-glucosidase
441	34	43.6	397	1	CBRZ	ubiquinol-cytochro	514	34	43.6	615	1	KPHU12	coagulation factor
442	34	43.6	397	2	T14263	ubiquinol-cytochro	515	34	43.6	616	2	T29234	hypothetical prote
443	34	43.6	398	1	A22931	ubiquinol-cytochro	516	34	43.6	621	2	T06717	hypothetical prote
444	34	43.6	403	2	H87569	hypothetical prote	517	34	43.6	621	2	B95897	probable cellulose
445	34	43.6	404	1	S25953	ubiquinol-cytochro	518	34	43.6	647	1	QYFPGM	phosphoenolpyruvat
446	34	43.6	406	1	JC4996	biphenyl dioxygena	519	34	43.6	656	2	C71361	probable long-chai
447	34	43.6	406	2	A37266	probable protease	520	34	43.6	661	2	T37753	hypothetical prote
448	34	43.6	406	2	B97108	collagenase family	521	34	43.6	676	1	A39379	hatching-suppress
449	34	43.6	408	2	C82977	glycolate oxidase	522	34	43.6	727	2	A55628	probable cellulos
450	34	43.6	415	2	T40338	ccaat-binding fact	523	34	43.6	762	2	G88436	protein T04A8.13 [
451	34	43.6	417	2	A43929	carboxypeptidase A	524	34	43.6	791	2	T24435	hypothetical prote
452	34	43.6	420	2	B87201	hypothetical prote	525	34	43.6	799	2	T02981	1,4-alpha-glucan b
453	34	43.6	423	2	E84027	molybdopterin bios	526	34	43.6	799	2	T01663	1,4-alpha-glucan b
454	34	43.6	425	2	G83046	probable acyl-CoA	527	34	43.6	807	2	A38152	F-spondin - rat
455	34	43.6	438	2	AD0437	conserved hypothet	528	34	43.6	809	2	E90016	hyaluronate lyase
456	34	43.6	444	2	G84565	probable flavonol	529	34	43.6	877	1	A25962	mitosis inhibitor
457	34	43.6	450	2	E98303	hypothetical 49.3K	530	34	43.6	879	2	F85875	probable fibrinial
458	34	43.6	450	2	AH2979	nitrotriacetate	531	34	43.6	879	2	E91031	probable outer mem
459	34	43.6	473	1	RXBCLT	ribulose-bisphosph	532	34	43.6	907	2	T27317	hypothetical prote
460	34	43.6	475	2	G81227	probable amino-aci	533	34	43.6	1022	2	S50534	hypothetical prote
461	34	43.6	477	2	S77373	hypothetical prote	534	34	43.6	1045	2	S23570	pol polyprotein ho
462	34	43.6	485	2	A82165	hypothetical prote	535	34	43.6	1138	2	H84529	hypothetical prote
463	34	43.6	486	2	B86411	protein R3M18.4 [i	536	34	43.6	1224	2	S25952	gene cob intron 3
464	34	43.6	494	2	G97496	integral membrane	537	34	43.6	1227	1	XYECMH	5-methyltetrahydro
465	34	43.6	497	2	AD3514	D-serine/D-alanine	538	34	43.6	1227	2	AH1011	5-methyltetrahydro
466	34	43.6	508	2	S74200	cytochrome-c oxida	539	34	43.6	1227	2	G86093	hypothetical prote
467	34	43.6	509	2	A71254	probable glu-tRNA	540	34	43.6	1227	2	A98246	hypothetical prote

541	34	43.6	1237	2	T45070	protein kinase hom	614	33	42.3	288	2	B70715	hypothetical prote
542	34	43.6	1300	2	I53799	CGI protein - huma	615	33	42.3	294	2	T10858	probable carboxype
543	34	43.6	1314	2	H86327	protein F18014.25	616	33	42.3	294	2	T38758	60s ribosomal prot
544	34	43.6	1356	2	S32763	kinectin 1 - human	617	33	42.3	294	2	T39325	60s ribosomal prot
545	34	43.6	1416	2	E88550	protein ZC94.1 [im	618	33	42.3	297	1	JN0773	calponin H1 - rat
546	34	43.6	1637	2	T00070	hypothetical prote	619	33	42.3	297	1	S31484	calponin H1 - pig
547	34	43.6	1734	2	A54602	microtubule-associ	620	33	42.3	297	1	S31486	calponin H1 - mous
548	34	43.6	1955	1	AGCH	agrin precursor -	621	33	42.3	297	1	JC4500	basic calponin - h
549	34	43.6	1974	2	T30010	hypothetical prote	622	33	42.3	297	2	G02142	smooth muscle cell
550	34	43.6	2233	1	ZLNZP3	genome polyprotein	623	33	42.3	301	2	C82394	probable glycine c
551	34	43.6	2844	2	S28291	hypothetical prote	624	33	42.3	301	2	G02142	probable ATP-bind
552	34	43.6	6260	2	T30228	polyketide synthas	625	33	42.3	308	1	H65050	hypothetical prote
553	34	43.6	8563	2	T30226	polyketide synthas	626	33	42.3	308	2	C85919	hypothetical prote
554	33.5	42.9	296	2	T06025	transcription fact	627	33	42.3	308	2	D91074	probable ATP-bind
555	33.5	42.9	310	2	T46035	ATMVB84 - Arabidop	628	33	42.3	308	2	J74834	hypothetical prote
556	33.5	42.9	321	1	E84026	probable S-adenosy	629	33	42.3	310	2	G86490	epidermal growth f
557	33.5	42.9	333	2	T51650	protein T30E16.21	630	33	42.3	310	2	T33205	probable biotin ho
558	33.5	42.9	348	2	B96620	myb-related protei	631	33	42.3	311	2	T43794	hypothetical prote
559	33.5	42.9	374	2	T05891	APX-1 protein homo	632	33	42.3	311	2	F90116	guanine nucleotide
560	33.5	42.9	376	2	T51673	hypothetical prote	633	33	42.3	311	2	T43794	probable phosphosu
561	33.5	42.9	379	2	T16213	beta-galactosidase	634	33	42.3	321	2	AB0845	hypothetical prote
562	33.5	42.9	701	2	T16607	hypothetical prote	635	33	42.3	321	2	F96588	probable NADPH2:qu
563	33.5	42.9	1036	2	T05687	MEGF5 protein - ra	636	33	42.3	326	2	B95941	prolamin box bindi
564	33.5	42.9	1151	2	H71347	genome polyprotein	637	33	42.3	326	2	T02046	hypothetical prote
565	33.5	42.9	1523	2	T13953	genome polyprotein	638	33	42.3	328	2	T09897	hypothetical prote
566	33.5	42.9	2430	1	GNWVW	genome polyprotein	639	33	42.3	329	2	T14760	hypothetical prote
567	33.5	42.9	3432	1	GNWVJ	genome polyprotein	640	33	42.3	331	2	T20593	hypothetical prote
568	33.5	42.9	3432	1	GNWVJ	genome polyprotein	641	33	42.3	331	2	S49054	hypothetical prote
569	33.5	42.9	3433	1	GNWVK	genome polyprotein	642	33	42.3	333	2	B42476	probable glycosylt
570	33	42.3	97	1	D44001	vpr protein - huma	643	33	42.3	333	2	AC0766	hypothetical prote
571	33	42.3	97	2	S51924	hypothetical prote	644	33	42.3	337	2	E89136	conserved hypothet
572	33	42.3	100	2	Q0859	hypothetical prote	645	33	42.3	339	2	H71361	hypothetical prote
573	33	42.3	101	2	S26314	IG heavy chain V r	646	33	42.3	340	2	T44739	hypothetical prote
574	33	42.3	109	2	AF0682	probable membrane	647	33	42.3	342	2	T31913	citrate synthase [
575	33	42.3	126	2	H70634	hypothetical prote	648	33	42.3	344	2	AD3197	hypothetical prote
576	33	42.3	153	2	D24372	restriction system	649	33	42.3	345	2	S72833	periplasmic bindin
577	33	42.3	153	2	F98079	DpnB protein (limp	650	33	42.3	347	2	H89835	hypothetical prote
578	33	42.3	153	2	H95215	DpnB protein (limp	651	33	42.3	349	2	H85626	hypothetical prote
579	33	42.3	157	2	B69659	molybdopterin bios	652	33	42.3	349	2	E90819	hypothetical prote
580	33	42.3	163	2	G90437	hypothetical prote	653	33	42.3	349	2	E90912	hypothetical prote
581	33	42.3	184	2	F69273	cytochrome-c oxida	654	33	42.3	367	2	S59146	ubiquinol-cytochro
582	33	42.3	193	1	JN0728	hypothetical prote	655	33	42.3	367	2	T52363	ubiquinol-cytochro
583	33	42.3	195	2	T46072	DNA binding protei	656	33	42.3	371	2	E72665	hypothetical prote
584	33	42.3	204	1	QXBP9L	hypothetical prote	657	33	42.3	372	2	A70302	rod shape determin
585	33	42.3	207	2	D90730	Ning protein (limp	658	33	42.3	372	2	AB3466	mandelate racemase
586	33	42.3	207	2	A85581	unknown protein en	659	33	42.3	377	2	T11361	ubiquinol-cytochro
587	33	42.3	210	2	T10647	hypothetical prote	660	33	42.3	378	2	B59180	wnt inhibitory fac
588	33	42.3	211	2	E70810	probable phosphate	661	33	42.3	380	2	G97828	hypothetical prote
589	33	42.3	221	2	G70533	phosphate transpor	662	33	42.3	381	2	S60623	ubiquinol-cytochro
590	33	42.3	222	2	G87182	hypothetical prote	663	33	42.3	382	2	T11338	ubiquinol-cytochro
591	33	42.3	224	2	T32606	hypothetical prote	664	33	42.3	383	2	S49928	F420-nonreducing h
592	33	42.3	226	2	G86641	hypothetical prote	665	33	42.3	383	2	S67477	F420-nonreducing h
593	33	42.3	228	2	T01350	hypothetical prote	666	33	42.3	385	1	S01511	ubiquinol-cytochro
594	33	42.3	230	2	C95028	phosphoglycerate m	667	33	42.3	385	2	C83611	ubiquinol-cytochro
595	33	42.3	231	2	E86743	phosphoglycerate m	668	33	42.3	399	2	B69531	ubiquinol-cytochro
596	33	42.3	232	2	T10646	hypothetical prote	669	33	42.3	403	2	T32607	ubiquinol-cytochro
597	33	42.3	232	2	T09898	hypothetical prote	670	33	42.3	409	2	A12823	ubiquinol-cytochro
598	33	42.3	233	2	T01867	hypothetical prote	671	33	42.3	414	2	H97601	ubiquinol-cytochro
599	33	42.3	239	2	AE2410	hypothetical prote	672	33	42.3	414	2	A12823	ubiquinol-cytochro
600	33	42.3	244	2	T31838	hypothetical prote	673	33	42.3	417	2	S07286	ubiquinol-cytochro
601	33	42.3	244	2	T00107	ABC-type transport	674	33	42.3	418	2	G83904	long-chain fatty a
602	33	42.3	247	2	T43382	ribosomal protein	675	33	42.3	428	2	B82248	hypothetical prote
603	33	42.3	248	2	C97899	phosphoglycerate m	676	33	42.3	432	2	B82248	hypothetical prote
604	33	42.3	249	2	T04939	hypothetical prote	677	33	42.3	432	2	A85040	GGDEF family prote
605	33	42.3	249	2	T21920	hypothetical prote	678	33	42.3	432	2	T01925	GGDEF family prote
606	33	42.3	251	2	AB2352	probable phosphor	679	33	42.3	443	2	F82231	hypothetical prote
607	33	42.3	252	2	C71430	hypothetical prote	680	33	42.3	446	2	T35094	cysteine-tRNA syn
608	33	42.3	259	2	E84694	probable tropinone	681	33	42.3	447	2	G95068	cysteine-tRNA liga
609	33	42.3	262	2	F89833	hypothetical prote	682	33	42.3	447	2	G97936	cysteine-tRNA liga
610	33	42.3	263	2	T33290	teichoic acid tran	683	33	42.3	450	2	T01711	probable serine/th
611	33	42.3	264	2	F89833	probable transposa	684	33	42.3	450	2	C86856	p-aminobenzoate sy
612	33	42.3	272	2	H95314	hypothetical prote	685	33	42.3	453	1	AGEC1	p-aminobenzoate sy
613	33	42.3	278	2	F84127	hypothetical prote	686	33	42.3	453	2	A99944	p-aminobenzoate sy
										454	2	A97048	similar to phospho
										457	2	D71717	cysteine-tRNA liga

687	33	42.3	461	2	A97491	cysteiny1-trNA syn	760	33	42.3	682	1	JH0560	cyclic nucleotide-
688	33	42.3	461	2	A2708	cysteiny1-trNA syn	761	33	42.3	707	2	T01502	probable serine/th
689	33	42.3	463	2	B56306	cysteiny1-trNA syn	762	33	42.3	713	2	A11752	endopeptidase [bac
690	33	42.3	463	2	B56306	dopamine receptor	763	33	42.3	716	1	BVECAD	membrane protein t
691	33	42.3	465	1	F64630	cysteine-trNA liga	764	33	42.3	738	2	B45046	basic juvenile hor
692	33	42.3	465	2	T16835	hypothetical prote	765	33	42.3	748	2	S45046	ABC transporter, A
693	33	42.3	465	2	D71884	cysteine-trNA liga	766	33	42.3	757	2	B75437	glycolate oxidase
694	33	42.3	466	2	C98820	cysteiny1-trNA syn	767	33	42.3	761	2	H65083	alpha-1 proteinase
695	33	42.3	469	2	B70607	probable cys prot	768	33	42.3	785	2	A29953	diacylglycerol kin
696	33	42.3	470	2	G86516	dicarboxylase tran	769	33	42.3	788	1	I59282	connectin 1 - chic
697	33	42.3	470	2	H72105	dicarboxylase tran	770	33	42.3	811	2	PN0689	connectin 1 - chic
698	33	42.3	471	2	G83561	hypothetical prote	771	33	42.3	828	2	S52393	beta-galactosidase
699	33	42.3	472	1	RKXR11	ribulose-bisphosph	772	33	42.3	839	2	C84685	probable beta-gala
700	33	42.3	473	2	S18315	ribulose-bisphosph	773	33	42.3	840	2	G72468	hypothetical prote
701	33	42.3	473	2	C86949	probable cysteinyl	774	33	42.3	849	2	G84533	hypothetical prote
702	33	42.3	476	2	D71108	cysteine-trNA liga	775	33	42.3	858	2	JC2309	chitin synthase (E
703	33	42.3	476	2	G84634	probable prolylcar	776	33	42.3	858	2	JC2308	chitin synthase (E
704	33	42.3	489	2	B69680	para-nitrobenzyl e	777	33	42.3	859	2	S62441	chitin synthase (E
705	33	42.3	491	2	S74473	probable starch sy	778	33	42.3	877	2	S49197	envelope protein p
706	33	42.3	493	2	A45737	phenylcarbamate hy	779	33	42.3	889	2	JC6015	chitin synthase (E
707	33	42.3	497	2	A86146	hypothetical prote	780	33	42.3	895	2	B64238	isoleucine-trNA li
708	33	42.3	501	2	T27613	hypothetical prote	781	33	42.3	904	2	A84212	hypothetical prote
709	33	42.3	503	2	C71535	probable NADH (ubi	782	33	42.3	908	2	T22376	hypothetical prote
710	33	42.3	506	2	AB3411	cysteine-trNA liga	783	33	42.3	910	2	JC4609	chitin synthase (E
711	33	42.3	507	1	C53308	probable hydro-lya	784	33	42.3	911	2	JC6016	chitin synthase (E
712	33	42.3	509	2	A31024	NADH dehydrogenase	785	33	42.3	916	2	JC2315	chitin synthase (E
713	33	42.3	509	2	C64999	NADH2 dehydrogenase	786	33	42.3	923	2	C97487	hypothetical prote
714	33	42.3	509	2	B85868	NADH dehydrogenase	787	33	42.3	923	2	A22705	conserved hypothet
715	33	42.3	520	2	F69470	dipeptide ABC tran	788	33	42.3	944	2	T47246	chitin synthase (E
716	33	42.3	523	2	JC7556	linoleoyl-CoA tran	789	33	42.3	955	2	B84022	hypothetical prote
717	33	42.3	532	2	A34329	60K esterase (EC 3	790	33	42.3	956	2	JH0826	glutamate ionotrop
718	33	42.3	534	2	H82244	sensor histidine k	791	33	42.3	956	2	JH0885	glutamate receptor
719	33	42.3	539	2	A29923	carboxylesterase (792	33	42.3	960	2	A41638	chitin synthase (E
720	33	42.3	540	2	A75250	carboxylesterase,	793	33	42.3	979	2	JH0589	glutamate receptor
721	33	42.3	540	2	A31584	carboxylesterase,	794	33	42.3	979	2	JH0592	glutamate receptor
722	33	42.3	543	2	S35047	mucin JUK7 - human	795	33	42.3	980	2	I57936	glutamate receptor
723	33	42.3	547	2	A83017	conserved hypothet	796	33	42.3	980	2	T27342	hypothetical prote
724	33	42.3	547	2	C98267	hypothetical prote	797	33	42.3	983	2	A59054	chitin synthase (E
725	33	42.3	549	2	JX0054	carboxylesterase (798	33	42.3	1005	2	A42265	alpha-mannosidase
726	33	42.3	554	1	S34067	carboxylesterase (799	33	42.3	1009	2	S20538	chitin synthase (E
727	33	42.3	554	2	A39060	carboxylesterase (800	33	42.3	1013	2	JC2314	chitin synthase (E
728	33	42.3	554	2	G75076	conserved hypothet	801	33	42.3	1013	2	T10659	probable serine/th
729	33	42.3	557	2	A47162	thiolesterase B (E	802	33	42.3	1036	2	B83466	probable RND efflu
730	33	42.3	559	1	JC5408	carboxylesterase (803	33	42.3	1049	1	S51784	toxin III - Actino
731	33	42.3	561	2	S71597	carboxylesterase (804	33	42.3	1131	2	A23944	chitin synthase (E
732	33	42.3	561	2	S47655	carboxylesterase (805	33	42.3	1217	2	C86159	hypothetical prote
733	33	42.3	561	2	JC2447	carboxylesterase (806	33	42.3	1231	2	C84716	hypothetical prote
734	33	42.3	561	2	S62788	carboxylesterase (807	33	42.3	1257	2	S22764	neurocan precursor
735	33	42.3	562	2	S27800	elastase precursor	808	33	42.3	1268	2	S52781	neurocan - mouse
736	33	42.3	562	2	A55281	carboxylesterase (809	33	42.3	1295	2	S60179	pol polyprotein ho
737	33	42.3	563	2	D84717	probable cysteinyl	810	33	42.3	1401	2	T39225	MAP kinase kinase
738	33	42.3	565	2	S10367	carboxylesterase (811	33	42.3	1421	2	T00333	hypothetical prote
739	33	42.3	566	2	S19307	carboxylesterase (812	33	42.3	1451	2	B41185	alpha-2 macroglobu
740	33	42.3	567	1	A41010	carboxylesterase (813	33	42.3	1464	2	JC5144	murinoglobulin pre
741	33	42.3	570	2	T32061	gut-specific carbo	814	33	42.3	1473	2	A20872	ovostatin precurs
742	33	42.3	571	2	T10232	hypothetical prote	815	33	42.3	1476	2	A41185	alpha-2 macroglobu
743	33	42.3	571	2	A10506	probable sulfatase	816	33	42.3	1477	2	A29952	alpha-1 proteinase
744	33	42.3	580	2	T18439	hypothetical prote	817	33	42.3	1487	2	S15904	alpha-1 proteinase
745	33	42.3	582	2	C69158	sensory transducti	818	33	42.3	1495	2	S27001	alpha-2-macroglobu
746	33	42.3	588	2	E87521	peptidase M1 fami1	819	33	42.3	1500	2	A42210	alpha-1-macroglobu
747	33	42.3	596	1	HQZJUL	hydrogenase (EC 1.	820	33	42.3	1574	2	T13954	MSGP6 protein - ra
748	33	42.3	596	1	S11969	hydrogenase (EC 1.	821	33	42.3	1743	2	T26859	hypothetical prote
749	33	42.3	600	2	T18446	hypothetical prote	822	33	42.3	1873	2	A30063	dihydropyridine re
750	33	42.3	601	1	S11777	hydrogenase (EC 1.	823	33	42.3	1963	1	MMKW	myosin heavy chain
751	33	42.3	602	1	JQ0806	hydrogenase (EC 1.	824	33	42.3	1992	1	S02771	319K protein ndvB
752	33	42.3	606	2	A70960	probable pckA prot	825	33	42.3	2870	2	A35548	319K protein ndvB
753	33	42.3	614	2	G71298	probable methyl-ac	826	33	42.3	3071	2	T50345	vacuolar protein s
754	33	42.3	617	2	T39660	hypothetical prote	827	33	42.3	4162	2	T42633	connectin/titin -
755	33	42.3	622	2	G81981	probable lipopolys	828	33	42.3	26926	1	I38344	titin, cardiac mus
756	33	42.3	644	2	C70585	hypothetical prote	829	32.5	41.7	80	2	H82194	conserved hypothet
757	33	42.3	667	2	T23010	hypothetical prote	830	32.5	41.7	154	2	T17816	hypothetical prote
758	33	42.3	670	2	AH2425	hypothetical prote	831	32.5	41.7	191	2	AC3413	holo-[acyl-carrier
759	33	42.3	681	2	S76354	ABC1-type transpor	832	32.5	41.7	224	2	T21788	hypothetical prote

833	32.5	41.7	297	2	B90754	hypothetical prote	106	32	41.0	222	2	A48935	25K protein - Rhiz
834	32.5	41.7	297	2	F64831	hypothetical prote	907	32	41.0	234	2	G97225	diverged arginase
835	32.5	41.7	297	2	H85617	hypothetical prote	908	32	41.0	234	2	AF0857	conserved hypothet
836	32.5	41.7	364	2	S74650	hypothetical prote	909	32	41.0	239	2	AB2704	ribonuclease III (
837	32.5	41.7	390	2	J4023	hypothetical prote	910	32	41.0	240	2	T33698	hypothetical prote
838	32.5	41.7	427	2	AG3499	transfoming growt	911	32	41.0	241	2	AP2773	transcription acti
839	32.5	41.7	427	2	A25928	aminopeptidase T (912	32	41.0	241	2	C97553	probable transcrip
840	32.5	41.7	453	2	A58375	cellulase (EC 3.2.	913	32	41.0	241	2	T23428	hypothetical prote
841	32.5	41.7	463	2	A48375	cellulase (EC 3.2.	914	32	41.0	246	2	D82078	conserved hypothet
842	32.5	41.7	464	2	JC7143	endoglucanase I -	915	32	41.0	248	2	S56531	hypothetical 27.4K
843	32.5	41.7	528	2	AG1810	glycogen (starch)	916	32	41.0	256	2	B97486	ribonuclease III (
844	32.5	41.7	528	2	T43919	yfufB protein [limp	917	32	41.0	259	2	G83131	conserved hypothet
845	32.5	41.7	623	2	H96766	unknown protein F2	918	32	41.0	260	2	S71507	restriction endonu
846	32.5	41.7	938	2	I49071	protein kinase - m	919	32	41.0	260	2	E87431	hypothetical prote
847	32.5	41.7	993	2	I48653	mouse developmen	920	32	41.0	261	2	G87535	regulatory protein
848	32.5	41.7	998	2	S37627	protein-tyrosine k	921	32	41.0	263	2	F84694	probable tropinone
849	32.5	41.7	1136	1	S57845	protein-tyrosine k	922	32	41.0	267	2	P97041	hypothetical prote
850	32.5	41.7	1525	1	GNWVS5	genome polyprotein	923	32	41.0	269	2	S73999	hypothetical prote
851	32.5	41.7	3988	1	GNWVBV	epidermal growth f	924	32	41.0	271	2	S38115	hypothetical prote
852	32.5	41.0	53	2	S17294	hypothetical prote	925	32	41.0	276	2	S57627	avirulence protein
853	32.5	41.0	59	2	E35523	hypothetical prote	926	32	41.0	278	2	E65020	ethanolamine utili
854	32.5	41.0	60	2	H97848	hypothetical prote	927	32	41.0	278	2	D91043	ethanolamine utili
855	32.5	41.0	67	2	T44097	hypothetical prote	928	32	41.0	278	2	G85887	probable ethanolam
856	32.5	41.0	71	2	AE3316	hypothetical prote	929	32	41.0	279	2	AC0814	cell surface glyco
857	32.5	41.0	75	2	B97578	hypothetical prote	930	32	41.0	281	1	B47629	hypothetical prote
858	32.5	41.0	95	2	T09383	vpr protein - huma	931	32	41.0	286	2	G70537	hypothetical prote
859	32.5	41.0	96	2	S36030	gene ace protein -	932	32	41.0	288	2	C70722	hypothetical prote
860	32.5	41.0	97	2	C82197	accessory cholera	933	32	41.0	288	2	T22721	H+-transporting tw
861	32.5	41.0	98	2	I40888	sarcosine oxidase	934	32	41.0	292	2	AE3283	probable DNA glyco
862	32.5	41.0	100	2	H72680	hypothetical prote	935	32	41.0	292	2	C87149	probable transcrip
863	32.5	41.0	101	2	G97567	hypothetical prote	936	32	41.0	302	2	D82996	fructokinase limpo
864	32.5	41.0	101	2	AP2788	hypothetical prote	937	32	41.0	308	2	AH2625	probable phosphati
865	32.5	41.0	108	2	S23204	retinol-binding pr	938	32	41.0	309	2	AB0343	hypothetical prote
866	32.5	41.0	109	2	S22204	hypothetical prote	939	32	41.0	310	2	E72667	fructokinase limpo
867	32.5	41.0	111	2	H95908	conserved hypothet	940	32	41.0	313	2	G97407	hypothetical prote
868	32.5	41.0	113	2	T42882	hypothetical 12.5K	941	32	41.0	316	2	T32962	transcription regu
869	32.5	41.0	114	1	QX13M	NADH2 dehydrogenas	942	32	41.0	319	2	AH2690	hypothetical trans
870	32.5	41.0	115	2	T11461	NADH2 dehydrogenas	943	32	41.0	319	2	D97472	tdTP-D-glucose 4,6
871	32.5	41.0	115	2	S68135	NADH2 dehydrogenas	944	32	41.0	328	2	AC1210	gonadotropin-relea
872	32.5	41.0	116	2	S55011	NADH2 dehydrogenas	945	32	41.0	328	2	JC1353	hypothetical prote
873	32.5	41.0	116	2	T11541	NADH2 dehydrogenas	946	32	41.0	328	2	T28099	hypothetical prote
874	32.5	41.0	116	2	T11771	NADH2 dehydrogenas	947	32	41.0	330	2	F82338	hypothetical prote
875	32.5	41.0	116	2	S08425	NADH2 dehydrogenas	948	32	41.0	336	2	F83225	hypothetical prote
876	32.5	41.0	116	2	T10993	NADH2 dehydrogenas	949	32	41.0	340	2	AE3310	coproporphyrinogen
877	32.5	41.0	116	2	H90626	NADH dehydrogenase	950	32	41.0	346	2	AF3161	phase-related inte
878	32.5	41.0	116	2	T11082	NADH2 dehydrogenas	951	32	41.0	355	2	I51309	major histocompati
879	32.5	41.0	116	2	T11294	NADH2 dehydrogenas	952	32	41.0	355	2	D86436	hypothetical prote
880	32.5	41.0	118	2	B0651	hypothetical prote	953	32	41.0	356	2	AB9612	protein F49E2.3 (i
881	32.5	41.0	119	2	S00227	phospholipase A2 (954	32	41.0	357	2	B56549	cell-cell signalin
882	32.5	41.0	126	2	D84283	hypothetical prote	955	32	41.0	359	2	A56549	wnt-5c protein - A
883	32.5	41.0	126	2	D69293	conserved hypothet	956	32	41.0	360	2	S34173	acetyltransferase,
884	32.5	41.0	136	2	G72573	hypothetical prote	957	32	41.0	363	2	B82090	probable RNA polym
885	32.5	41.0	137	2	AD3568	integral membrane	958	32	41.0	365	2	A48914	Wnt-5b protein - m
886	32.5	41.0	139	2	C96941	hypothetical prote	959	32	41.0	370	2	A70906	methylecitrate synt
887	32.5	41.0	144	2	D70392	hypothetical prote	960	32	41.0	372	2	E36470	ubiquinol-cytochro
888	32.5	41.0	154	2	E75024	hypothetical prote	961	32	41.0	377	2	AB2213	ubiquinol-cytochro
889	32.5	41.0	160	2	D96514	hypothetical prote	962	32	41.0	378	2	T12009	ubiquinol-cytochro
890	32.5	41.0	164	2	AD0813	ethanolamine utili	963	32	41.0	379	1	S26163	ubiquinol-cytochro
891	32.5	41.0	165	2	T20468	hypothetical prote	964	32	41.0	379	1	S41833	ubiquinol-cytochro
892	32.5	41.0	175	2	T40023	conserved hypothet	965	32	41.0	379	1	S41834	ubiquinol-cytochro
893	32.5	41.0	177	2	A45739	traf protein precu	966	32	41.0	379	1	S41847	ubiquinol-cytochro
894	32.5	41.0	184	2	A37041	homeotic protein A	967	32	41.0	379	1	S58461	ubiquinol-cytochro
895	32.5	41.0	187	2	C81131	hypothetical prote	968	32	41.0	379	2	T11152	ubiquinol-cytochro
896	32.5	41.0	192	2	C97818	hypothetical prote	969	32	41.0	379	2	S58465	ubiquinol-cytochro
897	32.5	41.0	198	2	S39543	GTP-binding protei	970	32	41.0	379	2	S58449	ubiquinol-cytochro
898	32.5	41.0	202	2	T04906	hypothetical prote	971	32	41.0	379	2	S58464	ubiquinol-cytochro
899	32.5	41.0	203	2	T30317	chemotaxis protein	972	32	41.0	379	2	S58447	ubiquinol-cytochro
900	32.5	41.0	211	2	B75470	hydrolase, CbbY/Cb	973	32	41.0	379	2	S58448	ubiquinol-cytochro
901	32.5	41.0	212	2	S28087	gene C protein - y	974	32	41.0	379	2	S58450	ubiquinol-cytochro
902	32.5	41.0	212	2	G95391	probable oxidoredu	975	32	41.0	379	2	D36470	Wnt-5a protein - m
903	32.5	41.0	216	2	D96946	transcription regu	976	32	41.0	380	1	CBXL	ubiquinol-cytochro
904	32.5	41.0	218	2	D83161	hypothetical prote	977	32	41.0	380	1	S36011	ubiquinol-cytochro
905	32.5	41.0	219	2	S27610	hypothetical prote	978	32	41.0	380	1		

Query Match 57.7%; Score 45; DB 2; Length 503;
 Best Local Similarity 77.8%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
 ||| |||||
 DB 379 LTEWPHGGF 387

RESULT 4
 CFBYY
 carboxypeptidase C (EC 3.4.16.5) precursor [validated] - yeast (Saccharomyces cerevisiae)
 N:Alternate names: carboxypeptidase Y; protein YMR297w
 C:Species: Saccharomyces cerevisiae
 C:Date: 15-Oct-1982 #sequence_revision 31-Mar-1993 #text_change 16-Aug-2004
 C:Accession: A26597; S47458; A90763; A94609; A00909
 R:Valls, L.A.; Hunter, C.P.; Rothman, J.H.; Stevens, T.H.
 Cell 48, 887-897, 1987
 A:Title: Protein sorting in yeast: the localization determinant of yeast vacuolar carbox
 A:Reference number: A26597; MUID:87131100; PMID:3028649
 A:Accession: A26597
 A:Molecule type: DNA
 A:Residues: 1-532 <VAL>
 A:Cross-references: UNIPROT:P00729; UNIPARC:UPI00001271DF; EMBL:M15482; NID:gl172238; PID
 R:Barrell, B.G.
 submitted to the EMBL Data Library, August 1994
 A:Reference number: S47445
 A:Accession: S47458
 A:Molecule type: DNA
 A:Residues: 1-532 <BAR>
 A:Cross-references: UNIPARC:UPI00001271DF; EMBL:X80836; NID:gl1289327; PIDN:CAA56806.1; H
 R:Svendsen, I.; Martin, B.M.; Viswanatha, T.; Johansen, J.T.
 Carlsberg Res. Commun. 47, 15-27, 1982
 A:Title: Amino acid sequence of carboxypeptidase Y. II. Peptides from enzymatic cleavage
 A:Reference number: A90763
 A:Accession: A90763
 A:Molecule type: Protein
 A:Residues: 112-223, 'X', 225; 228-239, 'X', 241-259, 'HG', 262-267, 'X', 269-388, 'E', 390-451, 'N'
 A:Cross-references: UNIPARC:UPI0000172A4A; UNIPARC:UPI0000172A4B
 A:Note: this sequence has been revised in reference A94609
 R:Svendsen, I.
 submitted to the Atlas, October 1982
 A:Reference number: A94609
 A:Contents: disulfide bond
 A:Accession: A94609
 A:Molecule type: protein
 A:Residues: 224-227 <SV2>
 A:Cross-references: UNIPARC:UPI0000172A4C
 A:Note: this is a revision to the sequence in reference A90763
 R:Martin, B.M.; Svendsen, I.; Viswanatha, T.; Johansen, J.T.
 Carlsberg Res. Commun. 47, 1-13, 1982
 A:Title: Amino acid sequence of carboxypeptidase Y. Peptides from cleavage with cyanogen
 A:Reference number: A90762
 A:Contents: annotation; experimental details
 R:Endrizzi, J.A.; Remington, S.J.
 submitted to the Brookhaven Protein Data Bank, March 1994
 A:Reference number: A52472; PDB:LYSC
 A:Contents: annotation; X-ray crystallography at 2.8 angstroms; disulfide bonds
 C:Genetics:
 A:Gene: SGD:PRC1; LBC1; MIPS:YMR297w
 A:Cross-references: SGD:S0004912; MIPS:YMR297w
 A:Map position: 13R
 C:Superfamily: Serine carboxypeptidase
 C:Keywords: glycoprotein; hydrolase; serine carboxypeptidase; yeast vacuole; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:1-111/Domain: propeptide #status predicted <PRO>
 F:112-532/Product: carboxypeptidase C #status experimental <MAT>
 F:124,198,279/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:167-409,304-318,328-344,373-379/Disulfide bonds: #status experimental
 F:257/Active site: Ser #status experimental
 F:449,508/Active site: Asp, His #status predicted

Query Match 57.7%; Score 45; DB 1; Length 532;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
 :||:|||||
 DB 522 VNEWINGGF 530

RESULT 5
 G94512
 hypotheical protein At2g13980 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: G84512
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: G84512
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-171 <STO>
 A:Cross-references: UNIPROT:Q9ZPU0; UNIPARC:UPI00000AA3DE; GB:AE002093; NID:g4388822; PI
 C:Genetics:
 A:Gene: At2g13980
 A:Map position: 2

Query Match 56.4%; Score 44; DB 2; Length 171;
 Best Local Similarity 58.3%; Pred. No. 5.4;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
 ||| |||||:
 DB 156 CTLPIWLHQHYC 167

RESULT 6
 AS3691
 diacylglycerol kinase (EC 2.7.1.107) gamma - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
 C:Accession: AS3691
 R:Kai, M.; Sakane, F.; Imai, S.; Wada, I.; Kanoh, H.
 J. Biol. Chem. 269, 18492-18498, 1994
 A:Title: Molecular cloning of a diacylglycerol kinase isozyme predominantly expressed in
 A:Reference number: AS3691; MUID:94308084; PMID:8034597
 A:Accession: AS3691
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-791 <KAI>
 A:Cross-references: UNIPROT:P49619; UNIPARC:UPI000012DD21; GB:D26135; NID:gs516757; PIDN
 C:Genetics:
 A:Gene: GDB:DAGK3
 A:Cross-references: GDB:377343
 A:Map position: 12q13.3-12q13.3
 C:Superfamily: human diacylglycerol kinase; calmodulin repeat homology; protein kinase
 C:Keywords: ATP; calcium binding; duplication; EF hand; phosphotransferase; zinc
 F:175-207/Domain: calmodulin repeat homology <EF1>
 F:220-252/Domain: calmodulin repeat homology <EF2>
 F:272-321/Domain: protein kinase C zinc-binding repeat homology <KZ1>
 F:337-383/Domain: protein kinase C zinc-binding repeat homology <KZ2>

Query Match 55.1%; Score 43; DB 2; Length 791;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLREWLHGG 10
 :||:|||||
 DB 241 SLQEWVHGG 249

```
RESULT 7
Tl1335
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Corvus frugilegus mitochond
C;Species: mitochondrion Corvus frugilegus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: Tl1335
R;Harlid, A.; Arnason, U.
Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999
A;Title: Analyses of mitochondrial DNA nest ratite birds within the Neognathae-supportin
A;Reference number: 217262
A;Accession: Tl1335
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-380 <HAR>
A;Cross-references: UNIPROT:O79386; UNIPARC:UPI0000091807; EMBL:Y18522; PIDN:CAA77206.1
C;Genetics:
A;Genome: mitochondrion
A;Note: cytb
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;12-340/Domain: cytochrome b6 homology <CYB>
F;12-210/Domain: cytochrome b6 homology <CB6>
F;229-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;84.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 53.8%; Score 42; DB 2; Length 380;
Best Local Similarity 80.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGQF 11
||| ||| |||
Db 160 TLVWNLWGQF 169

RESULT 8
D72073
conserved hypothetical protein CP0274 [imported] - Chlamydophila pneumoniae (strains CWL
N;Alternate names: ct383 hypothetical protein
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: D72073; E81594
R;Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: D72073
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <ARN>
A;Cross-references: UNIPROT:Q92871; UNIPARC:UPI00000D408B; GB:AE001633; GB:AE001363; NID
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10694935
A;Accession: E81594
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <REA>
A;Cross-references: UNIPARC:UPI00000D408B; GB:AE002188; GB:AE002161; NID:g7189198; PIDN:
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CPn0480; CP0274

Query Match 52.6%; Score 41; DB 2; Length 218;
Best Local Similarity 41.7%; Pred. No. 21;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGQFC 12
|:: |:: |||
```

```
Db 104 CSIVSWVFGGLC 115

RESULT 9
D86550
CT383 hypothetical protein [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86550
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ia
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: D86550
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <STO>
A;Cross-references: UNIPROT:Q92871; UNIPARC:UPI00000D408B; GB:BA000008; NID:g8978850; PI
A;Experimental source: strain J138
C;Genetics:
A;Gene: CPJ0480

Query Match 52.6%; Score 41; DB 2; Length 218;
Best Local Similarity 41.7%; Pred. No. 21;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTLREWLHGQFC 12
|:: |:: |||
Db 104 CSIVSWVFGGLC 115

RESULT 10
E96566
F6D8.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C;Accession: E96566
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96566
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <STO>
A;Cross-references: UNIPROT:Q95SR0; UNIPARC:UPI00000A4579; GB:AE005173; NID:gs903047; PI
C;Genetics:
A;Gene: F6D8.20
A;Map position: 1
C;Superfamily: AAA protein

Query Match 52.6%; Score 41; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 WLHGQF 11
|||||
Db 114 WLHGQF 119

RESULT 11
A81958
probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: A81958
```

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A;Reference number: AB1775; MUID:20222556; PMID:10761919
A;Accession: AB1958
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-530 <PAS>
A;Cross-references: UNIPROT:Q9JWE3; UNIPARC:UPI00000C497F; GB:AL162753; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0414

Query Match 52.6%; Score 41; DB 2; Length 530;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LREWLHGGGFC 12
||| ||| ||| |||
Db 195 LRPWLAGGVC 204

RESULT 12
EB1015
ABC transporter, permease protein NMB2026 [imported] - *Neisseria meningitidis* (strain MC
C;Species: *Neisseria meningitidis*
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
A;Accession: EB1015
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizze, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A>Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A;Reference number: AB1000; MUID:20175755; PMID:10710307
A;Accession: EB1015
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-531 <TET>
A;Cross-references: UNIPROT:Q9JXI9; UNIPARC:UPI00000C4849; GB:AE002552; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2026

Query Match 52.6%; Score 41; DB 2; Length 531;
Best Local Similarity 70.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LREWLHGGGFC 12
||| ||| ||| |||
Db 196 LRPWLAGGVC 205

RESULT 13
F82984
hypothetical protein PA5299 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C;Species: *Pseudomonas aeruginosa*
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A;Accession: F82984
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: F82984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <STO>
A;Cross-references: UNIPROT:Q9HTQ5; UNIPARC:UPI00000C5F94; GB:AE004942; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:

A;Gene: PA5299

Query Match 52.6%; Score 41; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WLHGGF 11
|||||
Db 329 WLHGGF 334

RESULT 14
H86250
hypothetical protein [imported] - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
A;Accession: H86250
R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
Chin, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A;Reference number: AB6141; MUID:21016719; PMID:11130712
A;Accession: H86250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1025 <STO>
A;Cross-references: UNIPROT:Q9SAA5; UNIPARC:UPI0000048234; GB:AE005172; NID:g4835785; PI
C;Genetics:
A;Map position: 1

Query Match 52.6%; Score 41; DB 2; Length 1025;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 WLHGGF 11
|||||
Db 332 WLHGGF 337

RESULT 15
A3443
pyrimidine synthesis multifunctional protein CAD - golden hamster
N;Contains: aspartate carbamoyltransferase (EC 2.1.3.2); carbamoyl-phosphate synthase (C
C;Species: *Mesocricetus auratus* (golden hamster)
C;Date: 29-Aug-1987 #sequence_revision 02-Jun-1994 #text_change 09-Jul-2004
A;Accession: A38653; A35432; PS0159; A23443; A30794; A34803; I48154
R;Bein, K.; Simmer, J.P.; Evans, D.R.
J. Biol. Chem. 266, 3791-3799, 1991
A>Title: Molecular cloning of a cDNA encoding the amino end of the mammalian multifuncti
A;Reference number: A38653; MUID:91139675; PMID:1671675
A;Accession: A38653
A;Molecule type: mRNA
A;Residues: 1-169 <BEI>
A;Cross-references: UNIPROT:P08955; UNIPARC:UPI0000170783; GB:M60078; NID:g191338; PIDN
R;Simmer, J.P.; Kelly, R.E.; Rinker Jr., A.G.; Scully, J.L.; Evans, D.R.
J. Biol. Chem. 265, 10395-10402, 1990
A>Title: Mammalian carbamyl phosphate synthetase (CPS). cDNA sequence and evolution of
A;Reference number: A35432; MUID:90285162; PMID:1972379
A;Accession: A35432
A;Molecule type: mRNA
A;Residues: 156-1455 <SIM>
A;Cross-references: UNIPARC:UPI0000170780; GB:J05503; NID:g191332; PIDN:AAA37062.1; PID
R;William, N.K.; Simpson, R.J.; Moritz, R.L.; Peide, Y.; Crofts, L.; Minasian, E.; Lea
Gene 94, 283-288, 1990
A>Title: Location of the dihydroorotase domain within trifunctional hamster dihydrooro
A;Reference number: PS0159; MUID:91078651; PMID:1979549
A;Accession: PS0159

A:Molecule type: mRNA
A:Residues: 1403-2110 <WIL>
A:Cross-references: UNIPARC:UPI000000E5DE8; GB:M33702; NID:g191172; PIDN:AAA37009.1; PID:
R/Shigeada, K.; Stark, G.R.; Maley, J.A.; Niswander, L.A.; Davidson, J.N.
Mol. Cell. Biol. 5, 1735-1742, 1985
A:Title: Construction of a cDNA to the hamster CAD gene and its application toward defin
A:Reference number: A23443; MUID:85267690; PMID:2862577
A:Accession: A23443
A:Molecule type: mRNA
A:Residues: 2074-2225 <SHI>
A:Cross-references: UNIPARC:UPI00001707A4; GB:M11242; NID:g191330; PIDN:AAA37061.1; PID:
R/Maley, J.A.; Davidson, J.N.
Biochem. Biophys. Res. Commun. 154, 1047-1053, 1988
A:Title: Identification of the junction between the glutamine amidotransferase and carba
A:Reference number: A30794; MUID:88309082; PMID:2900634
A:Accession: A30794
A:Molecule type: mRNA
A:Residues: 246-464 <MAL>
A:Cross-references: UNIPARC:UPI0000172F70; GB:M21927
R/Simmer, J.P.; Kelly, R.E.; Rinker Jr., A.G.; Zimmermann, B.H.; Scully, J.L.; Kim, H.;
Proc. Natl. Acad. Sci. U.S.A. 87, 174-178, 1990
A:Title: Mammalian dihydroorotase: nucleotide sequence, peptide sequences, and evolution
A:Reference number: A34803; MUID:90115834; PMID:1967494
A:Accession: A34803
A:Molecule type: mRNA
A:Residues: 1391-1870 <ST2>
A:Cross-references: UNIPARC:UPI0000170787; GB:M28866; NID:g191363; PIDN:AAA37073.1; PID:
A:Note: parts of this sequence were confirmed by peptide sequencing
R/Farnham, P.J.; Kollmar, R.
Cell Growth Differ. 1, 179-189, 1990
A:Title: Characterization of the 5' end of the growth-regulated Syrian hamster CAD gene.
A:Reference number: 148154; MUID:91190717; PMID:1982061
A:Accession: 148154
A:Molecule type: DNA
A:Residues: 1-73 <RES>
A:Cross-references: UNIPARC:UPI0000000486; GB:M31621; NID:g191326; PIDN:AAA37060.1; PID:
C/Genetics:
A:Gene: CAD
A:Map position: B9 short arm
A:Superfamily: 28/1
C:Superfamily: rudimentary enzyme; aspartate/ornithine carbamoyltransferase homology; Ba
arbanoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-phos
C/Keywords: hydrolase; ligase; methyltransferase; multifunctional enzyme; phosphoprotein
F/4-1442/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>
F/4-354/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain homolog
F/178-354/Domain: trpG homology <TRG>
F/395-1439/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain hom
F/395-845/Domain: biotin carboxylase homology <BCI>
F/934-1380/Domain: biotin carboxylase homology <BC2>
F/1457-1801/Domain: Bacillus dihydroorotase homology <DHO>
F/1924-2222/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>
F/252/Active site: Cys #status predicted
Query Match 52.6%; Score 41; DB 1; Length 2225;
Best Local Similarity 72.7%; Pred. NO. 2.2e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy 1 CTLREWL-HG 9
Db 102 CTLHEWLQHG 112
RESULT 16
H83343
probable non-ribosomal peptide synthetase PA2424 [imported] - Pseudomonas aeruginosa (st
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: H83343
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
J. Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10584043
A:Accession: H83343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4342 <STO>
A:Cross-references: UNIPROT:Q91157; UNIPARC:UPI0000110218; GB:AE004669; GB:AE004091; NID
A:Experimental source: strain PA01
C/Genetics:
A:Gene: PA2424
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F/60-553/Domain: acetate-CoA ligase homology <ACLI>
F/584-652/Domain: acyl carrier protein homology <ACP1>
F/1174-1622/Domain: acetate-CoA ligase homology <ACLI2>
F/1637-1705/Domain: acyl carrier protein homology <ACP2>
F/2332-2689/Domain: acetate-CoA ligase homology <ACLI3>
F/2706-2773/Domain: acyl carrier protein homology <ACP3>
F/3759-4230/Domain: acetate-CoA ligase homology <ACLI4>
F/4248-4316/Domain: acyl carrier protein homology <ACP4>
F/1669-2738,4280/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
Query Match 52.6%; Score 41; DB 2; Length 4342;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTLREWLH 8
Db 3559 CTVREWLJN 3566
RESULT 17
A38096
perlecan precursor - human
N/Alternate names: basement membrane heparan sulfate proteoglycan; heparan sulfate prote
C/Species: Homo sapiens (man)
C/Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 12-Jul-2004
C/Accession: A38096; S19256; S77946; A41059; A40306; B33625; A33625; A41736
R/Murdoch, A.D.; Dodge, G.R.; Cohen, I.; Tuan, R.S.; Iozzo, R.V.
J. Biol. Chem. 267, 8544-8557, 1992
A:Title: Primary structure of the human heparan sulfate proteoglycan from basement membr
tor, laminin, neural cell adhesion molecules, and epidermal growth factor.
A:Reference number: A38096; MUID:92235084; PMID:1569102
A:Accession: A38096
A:Molecule type: mRNA
A:Residues: 1-4391 <MUR>
A:Cross-references: UNIPROT:P98160; UNIPARC:UPI0000168756; GB:M85289; NID:g184426; PIDN:
R/Kallunki, P.; Tryggvason, K.
J. Cell Biol. 116, 559-571, 1992
A:Title: Human basement membrane heparan sulfate proteoglycan core protein: a 467-kD pro
ell adhesion molecules, and epidermal growth factor.
A:Reference number: A41736; MUID:92112994; PMID:11730768
A:Accession: S19256
A:Molecule type: mRNA
A:Residues: 1-57,'D',59-434,'A',436,'FL',438-449,'Q',451-502,'A',503-792,'K',794-908,'R',
71-2979,'H',2981-2994,'G',2996-3167,'T',3169-3240,'R',3242-3426,'R',3428-3631,'Q',3633-3-
A:Cross-references: UNIPARC:UPI000017CF3E; EMBL:X62515
R:Tryggvason, K.
submitted to the EMBL Data Library, October 1991
A:Reference number: S77946
A:Accession: S77946
A:Molecule type: mRNA
A:Residues: 1-57,'D',59-434,'A',436,'FL',438-449,'Q',451-502,'A',503-792,'K',794-908,'R',
71-2979,'H',2981-2994,'G',2996-3167,'T',3169-3240,'R',3242-3426,'R',3428-3631,'Q',3633-4-
A:Cross-references: UNIPARC:UPI000016866C; EMBL:X62515; NID:g29469; PIDN:CAA4373.1; PID
R/Kallunki, P.; Eddy, R.L.; Byers, M.G.; Keastilae, M.; Shows, T.B.; Tryggvason, K.
Genomics 11, 389-396, 1991
A:Title: Cloning of human heparan sulfate proteoglycan core protein, assignment of the g
A:Reference number: A41059; MUID:92120660; PMID:1685141
A:Accession: A41059
A:Molecule type: mRNA
A:Residues: 'RT',892-908,'R',910-1101,'L',1103-1132,'L',1134-1221,'L',1223-1397 <KA2>
A:Cross-references: UNIPARC:UPI000017C2E3; GB:S76436; NID:g243370; PIDN:AA821121.1; PID:
R/Dodge, G.R.; Kovalevsky, I.; Chu, M.L.; Hassell, J.R.; McBride, O.W.; Yi, H.F.; Iozzo,

Genomics 10, 673-680, 1991
A;Title: Heparan sulfate proteoglycan of human colon: partial molecular cloning, cellular
A;Reference number: A40306; MUID:91365376; PMID:1679749
A;Accession: A40306
A;Molecule type: mRNA
A;Residues: 1018-1405,'G',1407-1409,'G',1411-1465 <DOD>
A;Cross-references: UNIPARC:UPI000017C2E4; GB:M64283; NID:g184424; PID:AA52699.1; PID:
R;Heremans, A.; van der Schueren, B.; De Cock, B.; Paulsson, M.; Cassiman, J.J.; van den
J; Cell Biol. 109, 3199-3211, 1989
A;Title: Matrix-associated heparan sulfate proteoglycan: core protein-specific monoclonal
antibodies.
A;Reference number: A33625; MUID:90078352; PMID:2687294
A;Accession: B33625
A;Molecule type: protein
A;Residues: 1379-1384,'X',1386-1388,'X',1390-1398 <HE2>
A;Cross-references: UNIPARC:UPI000017C2E5
A;Accession: A33625
A;Molecule type: protein
A;Residues: 2166-2171,'X',2173-2175,'X',2177-2185 <HE3>
A;Cross-references: UNIPARC:UPI000017C2E5
A;Note: peptide potentially matches four different regions of sequence shown
C;Genetics:
A;Gene: GDB:HSPG2
A;Cross-references: GDB:126372; OMIM:142461
A;Map position: lp36.1-lp36.1
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein; heparan sulfate; transmembrane
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-4391/Product: perlecan #status predicted <MAT>
F;22-193/Domain: I <DOM1>
F;194-530/Domain: II <DOM2>
F;199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;531-1676/Domain: III <DOM3>
F;1159-1206/Domain: laminin-type EGF-like homology <LEG>
F;1563-1610/Domain: laminin-type EGF-like homology <EG7>
F;1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F;1677-3686/Domain: IV <DOM4>
F;2007-2034/Domain: transmembrane #status predicted <TRM>
F;3687-4391/Domain: V <DOM5>
F;3845-3880/Domain: EGF homology <EGF1>
F;3888-3921/Domain: EGF homology <EGF>
F;3953-4106/Domain: laminin G repeat homology <LG2>
F;4147-4175/Domain: EGF homology <EGF2>
F;4149-4151/Region: motor neuron attachment (L-R-E) motif
F;4299-4301/Region: motor neuron attachment (L-R-E) motif
F;65,71,76/Binding site: heparan sulfate (Ser) (covalent) #status predicted
F;89,554,1755,2121,3072,3105,3279,3780,3836,4068/Binding site: carbohydrate (Asn) (covalent)
F;2995,3933,4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 51.9%; Score 40.5; DB 2; Length 4391;
Best Local Similarity 69.2%; Pred. No. 5.3e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 CTLREWLHG 12
| | | | |
DB 4147 CQLREPLHGTC 4159

RESULT 18
B87354
hypothetical protein CC0845 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87354
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.;
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87354

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-117 <STO>
A;Cross-references: UNIPROT:Q9A9W6; UNIPARC:UPI00000C71B8; GB:AE005673; NID:g13422100; P:
C;Genetics:
A;Gene: CC0845

Query Match 51.3%; Score 40; DB 2; Length 117;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTLREWLHG 9
| | | | |
DB 38 CVLGHWLHG 46

RESULT 19
S25970
hypothetical protein 139 - liverwort (Marchantia polymorpha) mitochondrion
C;Species: mitochondrion Marchantia polymorpha
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 09-Jul-2004
C;Accession: S25970
R;Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.; Kan
J. Mol. Biol. 223, 1-7, 1992
A;Title: Gene organization deduced from the complete sequence of liverwort Marchantia
A;Reference number: S25941; MUID:92114051; PMID:1731062
A;Accession: S25970
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-139 <ODA>
A;Cross-references: UNIPROT:P38470; UNIPARC:UPI000013B8D1; EMBL:M68929; NID:g786182; P:
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
C;Genetics:
A;Genome: mitochondrion
C;Superfamily: liverwort mitochondrion hypothetical protein 139
C;Keywords: mitochondrion

Query Match 51.3%; Score 40; DB 2; Length 139;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 REWLHG 12
| | | | |
DB 78 REWFLG 86

RESULT 20
G84305
hypothetical protein mutY [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84305
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L.
A;Title: Genome sequence of Halobacterium species NRC-1
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-312 <STO>
A;Cross-references: UNIPROT:Q9HPQ6; UNIPARC:UPI00000638B8; GB:AE004437; NID:g10581009;
C;Genetics:
A;Gene: mutY

Query Match 51.3%; Score 40; DB 2; Length 312;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTLREWLH 8
| | | | |

Db 215 CPLREWCH 222

RESULT 21

S39007
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - guinea pig mitochondrion
C;Species: mitochondrion *Cavia porcellus* (guinea pig)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: S39007
R;Ma, D.P.; Zharkikh, A.; Graur, D.; Vandenberg, J.L.; Li, W.H.
J. Mol. Evol. 36, 327-334, 1993
A;Title: Structure and evolution of opossum, guinea pig, and porcupine cytochrome b gene
A;Reference number: S33572; MUID:93301932; PMID:8315653
A;Accession: S39007
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-378 <MAD>
A;Cross-references: UNIPROT:Q9TBF9; UNIPARC:UPI0000174C8F
C;Genetics:
A;Gene: cob
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;10-338/Domain: cytochrome b homology <CBH>
F;10-208/Domain: cytochrome b6 homology <CB6>
F;135-51/Domain: transmembrane #status predicted <TM1>
F;180-98/Domain: transmembrane #status predicted <TM2>
F;116-132/Domain: transmembrane #status predicted <TM3>
F;177-199/Domain: transmembrane #status predicted <TM4>
F;220-338/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-244/Domain: transmembrane #status predicted <TM5>
F;287-303/Domain: transmembrane #status predicted <TM6>
F;322-342/Domain: transmembrane #status predicted <TM7>
F;353-368/Domain: transmembrane #status predicted <TM8>
F;82,181/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;96,195/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 378;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGF 11
|||: |||
Db 158 TLVEWIGGF 167

RESULT 22

S17406
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Arabian camel mitochondrion
C;Species: mitochondrion *Camelus dromedarius* (Arabian camel)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17406
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17406
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24952; UNIPARC:UPI000012895E; EMBL:X56281; NID:gi12854; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIGGF 168

RESULT 24

S17410
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Grevy's zebra mitochondrion
C;Species: mitochondrion *Equus grevyi* (Grevy's zebra)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004

F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIGGF 168

RESULT 23

S17408
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - black rhinoceros mitochondrion
C;Species: mitochondrion *Diceros bicornis* (black rhinoceros)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17408
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17408
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24954; UNIPARC:UPI00001289A5; EMBL:X56283; NID:gi12903; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWIGGF 168

RESULT 24

S17410
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Grevy's zebra mitochondrion
C;Species: mitochondrion *Equus grevyi* (Grevy's zebra)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004

C:Accession: S17410
 R:Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
 J. Mol. Evol. 32, 128-144, 1991
 A:Title: Evolution of the cytochrome b gene of mammals.
 A:Reference number: S17405; MUID:91178817; PMID:1901092
 A:Accession: S17410
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <IRW>
 A:Cross-references: UNIPROT:P24956; UNIPARC:UPI00001289C7; EMBL:X56282; NID:gl2948; PIDN
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b6 homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
 Best Local Similarity 70.0%; Pred. No. 54;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TLREWLGGF 11
 |||||:
 Db 159 TLVEWIGGF 168

RESULT 25
 S17410
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - bridled dolphin mitochondr
 C:Species: mitochondrion Stenella attenuata (bridled dolphin)
 C:Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S17415
 R:Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
 J. Mol. Evol. 32, 128-144, 1991
 A:Title: Evolution of the cytochrome b gene of mammals.
 A:Reference number: S17405; MUID:91178817; PMID:1901092
 A:Accession: S17415
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <IRW>
 A:Cross-references: UNIPROT:P24961; UNIPARC:UPI000016D629; EMBL:X56294; NID:gl3495; PIDN
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b6 homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
 Best Local Similarity 70.0%; Pred. No. 54;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TLREWLGGF 11
 |||||:
 Db 159 TLVEWIGGF 168

RESULT 26
 S17417
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b, isolate 1B - pantropical sp
 C:Species: mitochondrion Stenella longirostris (pantropical spinner dolphin)
 C:Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S17417; S17416
 R:Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
 J. Mol. Evol. 32, 128-144, 1991
 A:Title: Evolution of the cytochrome b gene of mammals.
 A:Reference number: S17405; MUID:91178817; PMID:1901092
 A:Accession: S17417
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <IRW>
 A:Cross-references: UNIPROT:P24962; UNIPARC:UPI00000001E0; EMBL:X56293; NID:gl3628; PIDN
 A:Experimental source: isolate 1B
 A:Accession: S17416
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-59, 'T', 61-97, 'M', 99-265, 'P', 267-299, 'I', 301-326, 'V', 328-379 <IR2>
 A:Cross-references: UNIPARC:UPI0000128B12; EMBL:X56292; NID:gl3626; PIDN:CAA39739.1; PIL
 C:Experimental source: isolate 1A
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase compl
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rel
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b6 homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
 Best Local Similarity 70.0%; Pred. No. 54;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TLREWLGGF 11
 |||||:
 Db 159 TLVEWIGGF 168

RESULT 27
 S33572
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - southern African porcupin.

C;Species: mitochondrion Hyatrix africanae australis (southern African porcupine)
C;Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S33572
J;Ma, D.P.; Zharkikh, A.; Graur, D.; VandeBerg, J.L.; Li, W.H.
J. Mol. Evol. 36, 327-334, 1993
A;Title: Structure and evolution of opossum, guinea pig, and porcupine cytochrome b gene
A;Reference number: S33572; MUID:93301932; PMID:8315653
A;Accession: S33572
A;Molecule type: DNA
A;Cross-references: UNIPROT:Q04910; UNIPARC:UPI000016D69D; EMBL:X70674; NID:g14012; PIDN
A;Note: residue 1 and the corresponding nucleotide sequence are not shown
C;Genetics:
A;Gene: cob
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex
with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 TLREWLGGF 11
Db 159 TLVEWINGGF 168

RESULT 28
S41832
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - blue whale mitochondrion
C;Species: mitochondrion Balaenoptera musculus (blue whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S41832
J; Mol. Evol. 37, 312-322, 1993
A;Title: Comparison between the complete mtDNA sequences of the blue and the fin whale,
A;Reference number: S41820; MUID:94141932; PMID:8308901
A;Accession: S41832
A;Molecule type: DNA
A;Cross-references: UNIPROT:P41285; UNIPARC:UPI0000128949; EMBL:X72204; NID:g414126; PID
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 TLREWLGGF 11
Db 159 TLVEWINGGF 168

RESULT 28
S41832
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - blue whale mitochondrion
C;Species: mitochondrion Balaenoptera musculus (blue whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S41832
J; Mol. Evol. 37, 312-322, 1993
A;Title: Comparison between the complete mtDNA sequences of the blue and the fin whale,
A;Reference number: S41820; MUID:94141932; PMID:8308901
A;Accession: S41832
A;Molecule type: DNA
A;Cross-references: UNIPROT:P41285; UNIPARC:UPI0000128949; EMBL:X72204; NID:g414126; PID
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>

F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 TLREWLGGF 11
Db 159 TLVEWINGGF 168

RESULT 29
S43261
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - minke whale mitochondrion
C;Species: mitochondrion Balaenoptera acutorostrata (minke whale, lesser rorqual)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43261
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43261
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41280; UNIPARC:UPI0000128944; EMBL:X75753; NID:g457761; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;36-52/Domain: transmembrane #status predicted <CB6>
F;81-99/Domain: transmembrane #status predicted <TM1>
F;117-133/Domain: transmembrane #status predicted <TM2>
F;178-200/Domain: transmembrane #status predicted <TM3>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 2 TLREWLGGF 11
Db 159 TLVEWINGGF 168

RESULT 30
S43262
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaenoptera bonaerensis
C;Species: mitochondrion Balaenoptera bonaerensis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43262
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compari
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43262
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

```
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41281; UNIPARC:UPI0000128945; EMBL:X75581; NID:G457762; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGGF 11
Db 159 TLVEWIGGF 168

RESULT 31
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - sei whale mitochondrion
C;Species: mitochondrion Baleenoptera borealis (sei whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43263
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43263
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41282; UNIPARC:UPI0000128946; EMBL:X75582; NID:G457763; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGGF 11
Db 159 TLVEWIGGF 168

RESULT 32
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Bryde's whale mitochondrion
C;Species: mitochondrion Baleenoptera edeni (Bryde's whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43264
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43264
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41283; UNIPARC:UPI0000128947; EMBL:X75583; NID:G457766; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGGF 11
Db 159 TLVEWIGGF 168

RESULT 33
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena glacialis mitoch
```

```
Db 159 TLVEWIGGF 168

RESULT 32
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Bryde's whale mitochondrion
C;Species: mitochondrion Baleenoptera edeni (Bryde's whale)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S43264
R;Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A;Title: Relationship of baleen whales established by cytochrome b gene sequence compar
A;Reference number: S43261; MUID:94150700; PMID:8107866
A;Accession: S43264
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:P41284; UNIPARC:UPI0000171CB6; EMBL:X75587
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>

Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGGF 11
Db 159 TLVEWIGGF 168

RESULT 33
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena glacialis mitoch
```


Query Match 51.3%; Score 40; DB 1; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGHGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 37
E58889
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - duckbill platypus mitochondrion
C:Species: mitochondrion Ornithorhynchus anatinus (duckbill platypus)
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C:Accession: E58889
R:Janke, A.; Gemmell, N.J.; Feldmaier-Fuchs, G.; von Haeseler, A.; Paabo, S.
J. Mol. Evol. 42, 153-159, 1996
A:Title: The mitochondrial genome of a monotreme--the platypus (Ornithorhynchus anatinus)
A:Reference number: A58888; MUID:97077300; PMID:8919867
A:Accession: E58889
A:Status: nucleic acid sequence not shown; translation not shown; not compared with conc
A:Molecule type: DNA
A:Residues: 1-379 <JAN>
A:Cross-references: UNIPROT:Q36461; UNIPARC:UPI0000128A72; GB:X83427; NID:gl469249; PIDN
A:Note: submitted to GenBank/EMBL/DBJ December, 1994
C:Genetics:
A:Map position: FOR14186-15325
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; plastoquinol
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: transmembrane #status predicted <TM2>
F:36-52/Domain: transmembrane #status predicted <TM3>
F:81-99/Domain: transmembrane #status predicted <TM4>
F:117-133/Domain: transmembrane #status predicted <TM4>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-209/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGHGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 38
I48133
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl
C:Species: mitochondrion Dactylosomys boliviensis (American spiny rat)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I48133
R:da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodents
A:Reference number: A49605; MUID:94184505; PMID:8136924
A:Accession: I48133
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <RES>
A:Cross-references: UNIPROT:Q34301; UNIPARC:UPI0000095C92; GB:L23339; NID:g996082; PIDN:
C:Genetics:
A:Genome: mitochondrion

A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGHGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 39
I48132
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl
C:Species: mitochondrion Dactylosomys dactylinus (American spiny rat)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I48132
R:da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodents
A:Reference number: A49605; MUID:94184505; PMID:8136924
A:Accession: I48132
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <RES>
A:Cross-references: UNIPROT:Q34306; UNIPARC:UPI0000092E09; GB:L23335; NID:g995850; PIDN:
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLGHGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 40
I48180
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Isoth
C:Species: mitochondrion Isothrix bistriatus (American spiny rat)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I48180
R:da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodent
A:Reference number: A49605; MUID:94184505; PMID:8136924
A:Accession: I48180
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <RES>
A:Cross-references: UNIPROT:Q34814; UNIPARC:UPI0000094264; GB:L23355; NID:g995846; PIDN:
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion

A:Experimental source: kidney

C:Genetics:

A:Genome: mitochondrion

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos

F:11-339/Domain: cytochrome b6 homology <CB6>

F:11-209/Domain: cytochrome b6 homology <CB6>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;

Best Local Similarity 70.0%; Pred. No. 54;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGSF 11
|||: |||

Db 159 TLVEWINGGF 168

RESULT 43

Tl1869

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - horse mitochondrion

C:Species: mitochondrion Equus caballus (domestic horse)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: Tl1869

R:Xu, X.; Arnason, U.

Gene 146, 357-362, 1994

A:Title: The complete mitochondrial DNA sequence of the horse, Equus caballus: Extensive

A:Reference number: Z17369; MUID:95047450; PMID:7958969

A:Accession: Tl1869

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-379 <XU>

A:Cross-references: UNIPROT:P48665; UNIPARC:UPI00001289F7; EMBL:X79547; NID:9577571; PID

C:Genetics:

A:Genome: mitochondrion

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos

F:11-339/Domain: cytochrome b6 homology <CB6>

F:221-209/Domain: cytochrome b6 homology <CB6>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;

Best Local Similarity 70.0%; Pred. No. 54;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGSF 11
|||: |||

Db 159 TLVEWINGGF 168

RESULT 44

Tl1492

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rabbit mitochondrion

C:Species: mitochondrion Oryctolagus cuniculus (domestic rabbit)

C:Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004

C:Accession: Tl1492; S13823

R:Giessi, C.; Gullberg, A.; Arnason, U.

Genomics 50, 161-169, 1998

A:Title: The complete mitochondrial DNA sequence of the rabbit, Oryctolagus cuniculus.

A:Reference number: Z17275; MUID:98317530; PMID:9653643

A:Accession: Tl1492

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-379 <GIS>

A:Cross-references: UNIPROT:P34863; UNIPARC:UPI0000128AC6; EMBL:AJ001588; NID:g3293006;

A:Note: the termination resulting from transcript polyadenylation is shown

R:Mignotte, F.; Gueride, M.; Champagne, A.M.; Mounolou, J.C.

Eur. J. Biochem. 194, 561-571, 1990

A:Title: Direct repeats in the non-coding region of rabbit mitochondrial DNA. Involvement

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <REY>
A;Cross-references: UNIPROT:O63910; UNIPARC:UPI000009694F; EMBL:AJ001562; NID:g3127895;
C;Genetics:
A;Gene: cytb
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b6 homology <CVB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 45
S58990
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - earthworm (Lumbricus terre
C;Species: mitochondrion Lumbricus terrestris (common earthworm)
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C;Accession: S58990
R;Boore, J.L.; Brown, W.M.
Genetics 141, 305-319, 1995
A;Title: Complete sequence of the mitochondrial DNA of the annelid worm Lumbricus terre
A;Reference number: S58985; MUID:96042914; PMID:8536978
A;Accession: S58990
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <BOO>
A;Cross-references: UNIPROT:Q34945; UNIPARC:UPI000128A2D; EMBL:U24570; NID:g984290; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC4
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;12-340/Domain: cytochrome b6 homology <CBH>
F;12-210/Domain: cytochrome b6 homology <CB6>
F;222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 160 TLVEWINGGF 169

RESULT 46
T11401
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - fat dormouse mitochondrion
C;Species: mitochondrion Myoxus glis (fat dormouse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11401
R;Reyes, A.; Pesole, C.
Mol. Biol. Evol. 15, 499-505, 1998
A;Title: Complete mitochondrial DNA sequence of the fat dormouse, Glis glis: further eva
A;Reference number: Z17267; MUID:98242079; PMID:9580978
A;Accession: T11401

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <REY>
A;Cross-references: UNIPROT:O63910; UNIPARC:UPI000009694F; EMBL:AJ001562; NID:g3127895;
C;Genetics:
A;Gene: cytb
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F;11-339/Domain: cytochrome b6 homology <CVB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 47
A53077
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Sperophilus richardsonii
C;Species: mitochondrion Sperophilus richardsonii
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C;Accession: A53077
R;Thomas, W.K.; Martin, S.L.
Mol. Phylogenet. Evol. 2, 330-336, 1993
A;Title: A recent origin of marmots.
A;Reference number: A53077; MUID:94326000; PMID:8049781
A;Accession: A53077
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <RES>
A;Cross-references: UNIPROT:P49341; UNIPARC:UPI000128B0A; GB:S73150; NID:g639944; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 48
I48134
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Echim
C;Species: mitochondrion Echymys chrysurus (American spiny rat)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 31-Dec-2004
C;Accession: I48134
R;da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimvid rodent
A;Reference number: A49605; MUID:94184505; PMID:8136924
A;Accession: I48134
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-379 <RES>
A;Cross-references: UNIPROT:Q34420; UNIPARC:UPI0000096931; GB:L23341; NID:9995848; PIDN:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b homology; cytochrome b6 homology; plastoquinol-plastocyanin
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;1221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 2; Length 379;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLGGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 49
CBRT
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rat mitochondrion
C;Species: mitochondrion Rattus norvegicus (Norway rat)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C;Accession: A00154; S04759
R;Kojke, K.; Kobayashi, M.; Yaginuma, K.; Taira, M.; Yoshida, E.; Imai, M.
Gene 20, 177-185, 1982
A;Title: Nucleotide sequence and evolution of the rat mitochondrial cytochrome b gene
A;Reference number: A00154; MUID:83158755; PMID:6299885
A;Accession: A00154
A;Molecule type: DNA
A;Residues: 1-380 <KOI>
A;Cross-references: UNIPROT:P00159; UNIPARC:UPI000016D61E; GB:J01436; NID:G343168; PIDN:
A;Note: the authors translated the codon ATA for residue 42 as Ile, CAC for residue 54
96 as Phe. In another figure, the amino acids at all these positions except position 89,
translation in having 11-Glu
R;Gadalaria, G.; Pepe, G.; De Candia, G.; Quagliariello, C.; Sbisa, E.; Saccone, C.
J. Mol. Evol. 28, 497-516, 1989
A;Title: The complete nucleotide sequence of the Rattus norvegicus mitochondrial genome:
A;Reference number: S04747; MUID:89362487; PMID:2504926
A;Accession: S04759
A;Molecule type: DNA
A;Residues: 1-82,'Q',84-152,'I',154-380 <GAD>
A;Cross-references: UNIPARC:UPI000016390E; EMBL:X14848; NID:G854269; PIDN:CAA32966.1; PI
C;Genetics:
A;Gene: cob
A;Map position: 87-94
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-339/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 380;

Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLGGF 11
|||: |||
Db 159 TLVEWINGGF 168

RESULT 50
CBMS
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - mouse mitochondrion
C;Species: mitochondrion Mus musculus (house mouse)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C;Accession: A00153
R;Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
Cell 26, 167-180, 1981
A;Title: Sequence and gene organization of mouse mitochondrial DNA.
A;Reference number: A00153; MUID:82137051; PMID:7332926
A;Accession: A00153
A;Molecule type: DNA
A;Residues: 1-381 <BIB>
A;Cross-references: UNIPROT:P00158; UNIPARC:UPI00000008FF; GB:J01420; NID:G342520; PIDN:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-339/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 51.3%; Score 40; DB 1; Length 381;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLGGF 11
|||: |||
Db 159 TLVEWINGGF 168

Search completed: May 12, 2006, 10:51:22
Job time : 32.6641 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 84.9231 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-49
Perfect score: 78
Sequence: 1 CTTLEWLHGGPC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	60.3	97	1	P07471 bos taurus
2	47	60.3	128	2	O9GAB0 9BIVA
3	46	59.0	128	2	O9GAB2_MERCA
4	46	59.0	128	2	O9GAB1_9BIVA
5	46	59.0	279	2	O6H5Y4_ORYSA
6	46	59.0	462	2	Q75D31_ASHGO
7	46	59.0	767	2	Q4XV12_PLACH
8	46	59.0	1429	2	Q4YVE2_PLABE
9	46	59.0	2558	2	Q8IKF6_PLAF7
10	46	59.0	2584	2	Q7PDT6_PLAYO
11	45	57.7	126	2	Q8TJEO_METAC
12	45	57.7	211	2	Q9ZNB4_PORGI
13	45	57.7	211	2	Q7MUZ9_PORGI
14	45	57.7	361	2	Q9ZAN1_PORGI
15	45	57.7	361	2	Q9ZAD0_PORGI
16	45	57.7	361	2	Q7M7B5_PORGI
17	45	57.7	361	2	Q7M7E9_PORGI
18	45	57.7	396	2	Q67UT0_ORYSA
19	45	57.7	503	2	Q9HRG8_HALSA
20	45	57.7	532	1	CBPY_YEAST
21	44	56.4	154	2	Q7VZA4_PROMPT
22	44	56.4	171	2	Q9ZPU0_ARATH
23	44	56.4	539	2	Q6CBE1_YARLI
24	44	56.4	731	2	Q61MS4_CAEBR
25	44	56.4	1540	2	Q628B7_CAEBR
26	44	56.4	4856	2	Q61EJ2_CAEBR
27	43	55.1	158	2	Q7UG36_SYNXP
28	43	55.1	182	2	Q9B598_CHILA
29	43	55.1	328	2	Q5V5F2_HALMA
30	43	55.1	380	2	Q9B640_9RODE
31	43	55.1	440	2	Q8WF54_9BIVA

32	43	55.1	558	2	Q4ICS6_GIBZE	Q4ICS6 gibberella
33	43	55.1	583	2	Q8CY51_BIFLO	Q8CY51 bifidobacte
34	43	55.1	605	2	Q59GS6_HUMAN	Q59GS6 homo sapien
35	43	55.1	766	2	Q5FWG1_HUMAN	Q5FWG1 homo sapien
36	43	55.1	791	1	DKKG_HUMAN	P49619 homo sapien
37	43	55.1	884	2	O53JX1_ORYSA	Q53JX1 oryza sativ
38	42.5	54.5	104	2	Q9D245_MOUSE	Q9D245 m mus muscu
39	42	53.8	58	2	Q9T204_BPLC2	Q9T204 lactococcus
40	42	53.8	112	2	Q85A85_CORMC	Q85A85 corvus macr
41	42	53.8	121	2	Q8BG00_MOUSE	Q8BG00 m mus muscu
42	42	53.8	128	2	P92837_RHYME	P92837 ephehrus pe
43	42	53.8	134	2	Q7YAZ1_9SAUR	Q7YAZ1 colobosaura
44	42	53.8	144	2	Q9MMC3_GUIGU	Q9MMC3 guira guira
45	42	53.8	144	2	Q9MMC4_9AVES	Q9MMC4 crotophaga
46	42	53.8	193	2	Q4VTU6_9SAUR	Q4VTU6 gymnodactyl
47	42	53.8	196	2	Q7YB81_9SYLV	Q7YB81 alcippe mor
48	42	53.8	203	2	Q4VTU5_9SAUR	Q4VTU5 gymnodactyl
49	42	53.8	213	2	Q4VTU4_9SAUR	Q4VTU4 gymnodactyl
50	42	53.8	215	2	Q34157_CORMC	Q34157 corvus macr
51	42	53.8	215	2	Q34078_9SYLV	Q34078 cettia diph
52	42	53.8	215	2	Q34901_LEILU	Q34901 leiothrix l
53	42	53.8	221	2	Q8M2E3_9SAUR	Q8M2E3 pachydactyl
54	42	53.8	221	2	Q8M2D8_9SAUR	Q8M2D8 pachydactyl
55	42	53.8	222	2	Q7YB76_9SYLV	Q7YB76 alcippe poi
56	42	53.8	223	2	Q9TGR7_9SAUR	Q9TGR7 aspidosceli
57	42	53.8	226	2	Q7YB78_9SYLV	Q7YB78 alcippe per
58	42	53.8	226	2	Q7YAZ4_9SAUR	Q7YAZ4 cordylus na
59	42	53.8	228	2	Q9MHR8_9SAUR	Q9MHR8 hemidactyli
60	42	53.8	233	2	Q7YAZ0_9SAUR	Q7YAZ0 leposoma sc
61	42	53.8	233	2	Q7YAV2_9SAUR	Q7YAV2 cnemidophor
62	42	53.8	238	2	Q34569_GUIGU	Q34569 guira guira
63	42	53.8	238	2	Q4VTU3_9SAUR	Q4VTU3 gymnodactyl
64	42	53.8	249	2	Q21086_9RODE	Q21086 peromyscus
65	42	53.8	249	2	Q21085_9RODE	Q21085 peromyscus
66	42	53.8	252	2	Q7YB80_9SYLV	Q7YB80 alcippe mor
67	42	53.8	253	2	Q7YB77_9SYLV	Q7YB77 alcippe per
68	42	53.8	256	2	Q9TGQ6_9SAUR	Q9TGQ6 aspidosceli
69	42	53.8	258	2	Q9TGQ5_9SAUR	Q9TGQ5 aspidosceli
70	42	53.8	261	2	Q956Q5_9SAUR	Q956Q5 aspidosceli
71	42	53.8	261	2	Q9TGP4_9SAUR	Q9TGP4 aspidosceli
72	42	53.8	263	2	Q9TGQ7_9SAUR	Q9TGQ7 aspidosceli
73	42	53.8	263	2	Q9TGQ9_9SAUR	Q9TGQ9 aspidosceli
74	42	53.8	263	2	Q9TGP8_9SAUR	Q9TGP8 aspidosceli
75	42	53.8	264	2	Q9TGR8_9SAUR	Q9TGR8 aspidosceli
76	42	53.8	266	2	Q9TGR3_9SAUR	Q9TGR3 aspidosceli
77	42	53.8	266	2	Q9TGP9_9SAUR	Q9TGP9 aspidosceli
78	42	53.8	266	2	Q9TGQ2_9SAUR	Q9TGQ2 aspidosceli
79	42	53.8	266	2	Q9TGQ8_9SAUR	Q9TGQ8 aspidosceli
80	42	53.8	267	2	Q9TGR4_9SAUR	Q9TGR4 aspidosceli
81	42	53.8	267	2	Q9XKP3_9SAUR	Q9XKP3 aspidosceli
82	42	53.8	267	2	Q9TGQ4_9SAUR	Q9TGQ4 aspidosceli
83	42	53.8	267	2	Q9XKN6_9SAUR	Q9XKN6 aspidosceli
84	42	53.8	267	2	Q9TGR5_9SAUR	Q9TGR5 aspidosceli
85	42	53.8	268	2	Q9TGR0_9SAUR	Q9TGR0 aspidosceli
86	42	53.8	268	2	Q9XKP1_9SAUR	Q9XKP1 aspidosceli
87	42	53.8	268	2	Q956Q6_9SAUR	Q956Q6 aspidosceli
88	42	53.8	268	2	Q9TGQ3_9SAUR	Q9TGQ3 aspidosceli
89	42	53.8	268	2	Q9TGP6_9SAUR	Q9TGP6 aspidosceli
90	42	53.8	269	2	Q9TGQ0_9SAUR	Q9TGQ0 aspidosceli
91	42	53.8	269	2	Q9XKN5_9SAUR	Q9XKN5 aspidosceli
92	42	53.8	269	2	Q9TGR6_9SAUR	Q9TGR6 aspidosceli
93	42	53.8	270	2	Q9TGR1_9SAUR	Q9TGR1 aspidosceli
94	42	53.8	270	2	Q6UKX4_9SAUR	Q6UKX4 loliaemus p
95	42	53.8	271	2	Q9TGP7_9SAUR	Q9TGP7 aspidosceli
96	42	53.8	271	2	Q9XKP2_9SAUR	Q9XKP2 aspidosceli
97	42	53.8	271	2	Q9TGP3_9SAUR	Q9TGP3 aspidosceli
98	42	53.8	271	2	Q9XKN9_9SYLV	Q9XKN9 alcippe poi
99	42	53.8	271	2	Q8M4S1_9SYLV	Q8M4S1 alcippe poi
100	42	53.8	271	2	Q9TQOL_9SAUR	Q9TQOL aspidosceli
101	42	53.8	271	2	Q9XKN8_9SAUR	Q9XKN8 aspidosceli
102	42	53.8	271	2	Q9XKN7_9SAUR	Q9XKN7 aspidosceli
103	42	53.8	271	2	Q9XKP0_9SAUR	Q9XKP0 aspidosceli
104	42	53.8	271	2	Q9TGR2_9SAUR	Q9TGR2 aspidosceli

105 42 53.8 271 2 Q9XK04_9SAUR
 106 42 53.8 272 2 Q7YB92_9SYLV
 107 42 53.8 276 2 Q8HD11_9PASS
 108 42 53.8 278 2 Q9GB00_9CORV
 109 42 53.8 278 2 Q9GB09_9CORV
 110 42 53.8 278 2 Q9GB06_9CORV
 111 42 53.8 278 2 Q9GB07_9CORV
 112 42 53.8 281 2 Q7YB82_9SYLV
 113 42 53.8 283 2 Q9B7J9_9EMBE
 114 42 53.8 283 2 Q9B7J6_9EMBE
 115 42 53.8 283 2 Q9B7K2_9EMBE
 116 42 53.8 283 2 Q9B7J8_9EMBE
 117 42 53.8 283 2 Q9B7K1_POOGA
 118 42 53.8 283 2 Q9B7J7_9EMBE
 119 42 53.8 283 2 Q9B7J5_9EMBE
 120 42 53.8 283 2 Q9B7K3_9EMBE
 121 42 53.8 283 2 Q9B7J0_9EMBE
 122 42 53.8 283 2 Q9B7J1_9EMBE
 123 42 53.8 283 2 Q9B1N1_POOAL
 124 42 53.8 284 2 Q9TG05_9CORV
 125 42 53.8 287 2 Q9WZ95_LEGPL
 126 42 53.8 287 2 Q9X7Z4_LEGPA
 127 42 53.8 287 2 Q9ZYH9_LEGPH
 128 42 53.8 287 2 Q9TFV5_LEILU
 129 42 53.8 287 2 Q9TFV4_GARLE
 130 42 53.8 287 2 Q9TG08_CORCE
 131 42 53.8 287 2 Q9TG03_9PASS
 132 42 53.8 293 2 Q6XK02_CORCE
 133 42 53.8 294 2 Q7YB79_9SYLV
 134 42 53.8 296 2 Q94Y05_9SYLV
 135 42 53.8 303 2 Q6LE47_9SYLV
 136 42 53.8 308 1 CYB_CORCX
 137 42 53.8 308 2 Q95629_9META
 138 42 53.8 308 2 Q92675_NUCCA
 139 42 53.8 308 2 Q92490_CVACH
 140 42 53.8 308 2 Q92608_CORMO
 141 42 53.8 308 2 Q92636_GARGL
 142 42 53.8 308 2 Q92637_GARLI
 143 42 53.8 308 2 Q92710_PERIN
 144 42 53.8 308 2 Q03813_PICPC
 145 42 53.8 308 2 Q92688_PTIAF
 146 42 53.8 308 2 Q92601_CORCE
 147 42 53.8 308 2 Q9WRH1_9CORV
 148 42 53.8 308 2 Q9WRG9_9CORV
 149 42 53.8 308 2 Q9WRG7_9CORV
 150 42 53.8 308 2 Q9WRG6_9CORV
 151 42 53.8 308 2 Q9WRG5_9CORV
 152 42 53.8 308 2 Q9WRG3_9CORV
 153 42 53.8 308 2 Q9DRL7_CORAL
 154 42 53.8 308 2 Q9DRL6_9CORV
 155 42 53.8 308 2 Q9DRU5_CORCY
 156 42 53.8 308 2 Q9DRU4_CORCY
 157 42 53.8 308 2 Q9DRL3_CORCX
 158 42 53.8 308 2 Q9DRL2_CORCX
 159 42 53.8 308 2 Q9DRK8_CORBR
 160 42 53.8 308 2 Q9DRU0_CORCX
 161 42 53.8 308 2 Q9SVJ8_9PASE
 162 42 53.8 331 2 Q6NY49_BRARE
 163 42 53.8 340 2 Q7Y9K4_9SYLV
 164 42 53.8 345 2 Q9VEB5_CORCX
 165 42 53.8 345 2 Q9YED4_CORCX
 166 42 53.8 345 2 Q9YEC8_CORCX
 167 42 53.8 345 2 Q9YEF3_CORCX
 168 42 53.8 345 2 Q9YEE4_CORCX
 169 42 53.8 345 2 Q9YEC5_CORCX
 170 42 53.8 346 2 Q70VV3_9SYLV
 171 42 53.8 347 2 Q63411_9EMBE
 172 42 53.8 347 2 Q63434_9EMBE
 173 42 53.8 347 2 Q9XMF4_9SYLV
 174 42 53.8 347 2 Q63428_9EMBE
 175 42 53.8 347 2 Q9G745_9SYLV
 176 42 53.8 347 2 Q63407_9EMBE
 177 42 53.8 347 2 Q63419_9EMBE

178 42 53.8 347 2 Q63440_9EMBE
 179 42 53.8 351 2 Q9MM82_9SAUR
 180 42 53.8 367 2 Q6Q0B4_APLCA
 181 42 53.8 369 2 Q8SJ48_9DINO
 182 42 53.8 373 2 Q8M4X4_9DINO
 183 42 53.8 380 1 CYB_APHCE
 184 42 53.8 380 1 CYB_CALFO
 185 42 53.8 380 1 CYB_CORBR
 186 42 53.8 380 1 CYB_CORCD
 187 42 53.8 380 1 CYB_CORFR
 188 42 53.8 380 1 CYB_CYACH
 189 42 53.8 380 1 CYB_CYAST
 190 42 53.8 380 2 Q8M4D3_9EMBE
 191 42 53.8 380 2 Q8M4D4_9EMBE
 192 42 53.8 380 2 P92622_CVAVI
 193 42 53.8 380 2 Q950C5_PICHU
 194 42 53.8 380 2 P92708_PERCN
 195 42 53.8 380 2 Q34066_CYACR
 196 42 53.8 380 2 Q950C4_APHCA
 197 42 53.8 380 2 Q8HJR8_9CORV
 198 42 53.8 380 2 Q7YB87_9SYLV
 199 42 53.8 380 2 Q7YB86_9SYLV
 200 42 53.8 380 2 Q7YB85_9SYLV
 201 42 53.8 380 2 Q7YB84_9SYLV
 202 42 53.8 380 2 Q7YB83_9SYLV
 203 42 53.8 380 2 Q7YB81_9SYLV
 204 42 53.8 380 2 Q7YB80_9SYLV
 205 42 53.8 380 2 Q85UG1_CORMC
 206 42 53.8 380 2 Q85UG0_CORMC
 207 42 53.8 380 2 Q85UF9_CORMC
 208 42 53.8 380 2 Q85UF8_CORMC
 209 42 53.8 380 2 Q85UF7_CORMC
 210 42 53.8 380 2 Q6WRE0_PAVES
 211 42 53.8 380 2 Q6I7W2_9SAUR
 212 42 53.8 380 2 Q85AB8_CORMC
 213 42 53.8 380 2 Q7Y888_9SYLV
 214 42 53.8 380 2 Q7Y889_9SYLV
 215 42 53.8 380 2 Q85AA7_CORMC
 216 42 53.8 380 2 Q85BT3_CORMC
 217 42 53.8 380 2 Q7Y848_9SYLV
 218 42 53.8 380 2 Q7Y614_9SYLV
 219 42 53.8 380 2 Q61892_9SYLV
 220 42 53.8 380 2 Q61875_9SYLV
 221 42 53.8 380 2 P92635_GYMCV
 222 42 53.8 380 2 Q7Y6B3_9SYLV
 223 42 53.8 380 2 Q61872_9SYLV
 224 42 53.8 381 1 CYB_ANTBE
 225 42 53.8 381 1 CYB_ANTFL
 226 42 53.8 381 1 CYB_ANTLE
 227 42 53.8 381 1 CYB_ANTMI
 228 42 53.8 381 1 CYB_ANTNA
 229 42 53.8 381 1 CYB_ANTST
 230 42 53.8 381 1 CYB_DASAL
 231 42 53.8 381 1 CYB_DASBY
 232 42 53.8 381 1 CYB_DASGE
 233 42 53.8 381 1 CYB_DASHA
 234 42 53.8 381 1 CYB_DASMA
 235 42 53.8 381 1 CYB_DASRO
 236 42 53.8 381 1 CYB_DASSP
 237 42 53.8 381 1 CYB_DASVI
 238 42 53.8 381 1 CYB_MURLO
 239 42 53.8 381 1 CYB_MURRO
 240 42 53.8 381 1 CYB_MYOBE
 241 42 53.8 381 1 CYB_MYOYA
 242 42 53.8 381 1 CYB_NEOLO
 243 42 53.8 381 1 CYB_NINRI
 244 42 53.8 381 1 CYB_NINTV
 245 42 53.8 381 1 CYB_NINTV
 246 42 53.8 381 1 CYB_PARPAP
 247 42 53.8 381 1 CYB_PAREI
 248 42 53.8 381 1 CYB_PHACL
 249 42 53.8 381 1 CYB_PHATA
 250 42 53.8 381 1

O63440 thlypopsis
 Q9mm82 neusticurus
 Q6q0b4 aphysia cal
 Q8sj48 pfiesteria
 Q8m4x4 pfiesteria
 P92493 aphelocoma
 P92605 calocitta f
 Q950C7 corvus brac
 Q950C8 corvus cor
 Q950C9 corvus frug
 P92599 cyanocorax
 Q950C6 cyanocitta
 Q8m4D3 hemispingus
 Q8m4D4 hemispingus
 P92622 cyanolyca v
 Q950C5 pica hudson
 P92708 parisoreus
 Q34066 cyanocitta
 Q950C4 aphelocoma
 Q8hjr8 terpsiphone
 Q7ybb7 garulax ca
 Q7ybb6 garulax ca
 Q7ybb5 garulax ca
 Q7ybb4 garulax po
 Q7ybb3 garulax ch
 Q7ybb1 garulax po
 Q7ybb0 garulax po
 Q85ug1 corvus macr
 Q85ug0 corvus macr
 Q85uf9 corvus macr
 Q85uf8 corvus macr
 Q85uf7 corvus macr
 Q6wre0 crotophaga
 Q6i7w2 cordylus wa
 Q85ab8 corvus macr
 Q7y888 garulax ca
 Q7y889 garulax ca
 Q85aa7 corvus macr
 Q85bt3 corvus macr
 Q7y848 garulax ca
 Q7y614 garulax ca
 Q61892 urophena s
 Q61875 cettia diph
 P92635 gymnorhinus
 Q7y6b3 garulax ca
 Q61872 cettia diph
 O63535 antechinus
 O32043 antechinus
 O32044 antechinus
 O63537 antechinus
 O63534 antechinus
 Q33800 antechinus
 P92509 antechinus
 Q34289 dasyurus al
 Q34300 dasyuroides
 O20604 dasyurus ge
 Q34321 dasyurus ha
 O3522 dasyurus ma
 Q34376 dasykaluta
 Q34382 dasyurus sp
 Q34399 dasyurus vi
 Q35020 murexia lon
 Q35078 murexia rot
 Q35038 myoictis me
 Q35157 neophascoga
 Q35172 ningau rid
 Q35173 ningau tim
 Q35196 ningau yro
 Q35377 parantechin
 Q35380 parantechin
 Q35409 phascogale
 Q35673 phascogale

251	42	53.8	381	1	CYB_PLAGI	Q35459 pianigale g	324	41	52.6	379	2	Q9T612_LEPSI	Q9t612 lepus sinen
252	42	53.8	381	1	CYB_PLAIN	Q35485 pianigale i	325	41	52.6	380	1	CYB_THOIS	Q9xw9 thomatomys
253	42	53.8	381	1	CYB_PLATE	Q35675 pianigale t	326	41	52.6	381	1	CYB_ANTME	Q33782 antechinus
254	42	53.8	381	1	CYB_PSEMD	Q03543 pseudantech	327	41	52.6	381	1	CYB_THYCV	Q36012 thylacinus
255	42	53.8	381	1	CYB_PSEMO	Q35695 pseudantech	328	41	52.6	381	2	Q9XP74_SMIVI	Q9xp74 sminthopsis
256	42	53.8	381	1	CYB_SARHA	Q35861 sarcophilus	329	41	52.6	391	2	Q4KAE2_PSEFS	Q4kae2 pseudomonas
257	42	53.8	381	1	CYB_SMAI	Q9xp89 sminthopsis	330	41	52.6	391	2	Q4SIH1_TETNG	Q4sih1 tetraodon n
258	42	53.8	381	1	CYB_SMIAR	Q9xp88 sminthopsis	331	41	52.6	460	2	Q7PR34_ANOGA	Q7pr34 anopheles g
259	42	53.8	381	1	CYB_SMICR	Q35810 sminthopsis	332	41	52.6	494	2	Q60VG8_CAEBR	Q60vg8 caenorhabdi
260	42	53.8	381	1	CYB_SMICR	Q9xp83 sminthopsis	333	41	52.6	530	2	Q5F558_NEIGI	Q5f558 neisseria g
261	42	53.8	381	1	CYB_SMIGA	Q20433 sminthopsis	334	41	52.6	530	2	Q5JWE3_NEIMA	Q5jwe3 neisseria m
262	42	53.8	381	1	CYB_SMIMA	Q35886 sminthopsis	335	41	52.6	531	2	Q9JX19_NEIMB	Q9jx19 neisseria m
263	42	53.8	381	1	Q9XP80_SMILE	Q9xp80 sminthopsis	336	41	52.6	542	2	Q8PF22_XANAC	Q8pf22 xanthomonas
264	42	53.8	381	2	Q9XP86_SMIDL	Q9xp86 sminthopsis	337	41	52.6	567	2	Q59H79_HUMAN	Q59h79 homo sapien
265	42	53.8	381	2	Q9XP87_SMIBI	Q9xp87 sminthopsis	338	41	52.6	619	2	Q9HTQ5_PSEAE	Q9htq5 pseudomonas
266	42	53.8	381	2	Q9XP82_SMIGR	Q9xp82 sminthopsis	339	41	52.6	619	2	Q4Q2C6_LEIMA	Q4q2c6 leishmania
267	42	53.8	381	2	Q9XP81_SMIMI	Q9xp81 sminthopsis	340	41	52.6	658	2	Q4Q2C6_LEIMA	Q4q2c6 leishmania
268	42	53.8	381	2	Q9XP78_SMILE	Q9xp78 sminthopsis	341	41	52.6	735	2	Q7YS57_BOVIN	Q7ys57 bos taurus
269	42	53.8	381	2	Q9XP79_SMILE	Q9xp79 sminthopsis	342	41	52.6	735	2	Q7YS57_BOVIN	Q7ys57 bos taurus
270	42	53.8	381	2	Q9XP81_SMIMI	Q9xp81 sminthopsis	343	41	52.6	763	2	Q7YS58_BOVIN	Q7ys58 bos taurus
271	42	53.8	381	2	Q35545_9META	Q35545 pianigale s	344	41	52.6	763	2	Q7YS58_BOVIN	Q7ys58 bos taurus
272	42	53.8	381	2	Q35545_9META	Q9xp76 sminthopsis	345	41	52.6	763	2	Q7YS58_BOVIN	Q7ys58 bos taurus
273	42	53.8	381	2	Q9XP76_SMIPS	Q9xp76 sminthopsis	346	41	52.6	785	2	Q7YS56_BOVIN	Q7ys56 bos taurus
274	42	53.8	381	2	Q9XP78_SMILE	Q9xp78 sminthopsis	347	41	52.6	785	2	Q7YS56_BOVIN	Q7ys56 bos taurus
275	42	53.8	381	2	Q9XP79_SMILE	Q9xp79 sminthopsis	348	41	52.6	796	2	Q541B5_MOUSE	Q541e5 mus musculu
276	42	53.8	381	2	Q63533_ANTST	Q63533 antechinus	349	41	52.6	826	2	Q7YS59_BOVIN	Q7ys59 bos taurus
277	42	53.8	381	2	Q33723_ANTCHIN	Q33723 antechinus	350	41	52.6	826	2	Q7YS59_BOVIN	Q7ys59 bos taurus
278	42	53.8	381	2	Q33723_ANTCHIN	Q9xp75 sminthopsis	351	41	52.6	853	1	DNM3B_HUMAN	Q9ubc3 homo sapien
279	42	53.8	381	2	Q5Q80_SMIDO	Q5q80 sminthopsis	352	41	52.6	853	1	DNM3B_HUMAN	Q9ubc3 homo sapien
280	42	53.8	381	2	Q5Q80_SMIDO	Q5q80 sminthopsis	353	41	52.6	859	1	DNM3B_MOUSE	Q88509 mus musculu
281	42	53.8	381	2	Q5J1T9_PHADO	Q5j1t9 phascosolor	354	41	52.6	869	2	Q6PI72_MOUSE	Q6pi72 mus musculu
282	42	53.8	381	2	Q5J1T8_PHADO	Q5j1t8 phascosolor	355	41	52.6	869	2	Q6PI72_MOUSE	Q6pi72 mus musculu
283	42	53.8	399	2	Q38192_BPLC2	Q38192 lactococcus	356	41	52.6	949	2	Q9ERT8_MOUSE	Q9ert8 mus musculu
284	42	53.8	415	2	Q38445_9VIRU	Q38445 bacteriococcus	357	41	52.6	949	2	Q9ERT8_MOUSE	Q9ert8 mus musculu
285	42	53.8	415	2	Q38445_9VIRU	Q38445 bacteriococcus	358	41	52.6	971	2	Q4QGF6_LEIMA	Q4qgf6 leishmania
286	42	53.8	416	2	Q53CH8_9CAUD	Q53ch8 lactococcus	359	41	52.6	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
287	42	53.8	416	2	Q53CH8_9CAUD	Q53ch8 lactococcus	360	40.5	51.9	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
288	42	53.8	416	2	Q53C12_BPLC2	Q53c12 lactococcus	361	40.5	51.9	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
289	42	53.8	692	2	Q5STZ6_CRYNE	Q5stz6 cryptococcus	362	40.5	51.9	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
290	42	53.8	692	2	Q5STZ6_CRYNE	Q5stz6 cryptococcus	363	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
291	42	53.8	966	2	Q5GA19_9VIRU	Q5ga19 groupier iri	364	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
292	42	53.8	966	2	Q5GA19_9VIRU	Q5ga19 groupier iri	365	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
293	42	53.8	968	2	Q5VFL3_SINGAPORE	Q5vfl3 singapore g	366	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
294	42	53.8	1179	2	Q53K95_ORYSA	Q53k95 oryza sativ	367	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
295	41.5	53.8	1200	2	Q9VPH4_DROME	Q9vph4 drosophila	368	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
296	41	52.6	107	2	Q7X9R3_GOSBA	Q7x9r3 gossypium b	369	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
297	41	52.6	75	2	Q8JZL6_MOUSE	Q8jzl6 mus musculu	370	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
298	41	52.6	103	2	Q7K6V2_CABEL	Q7k6v2 caenorhabdi	371	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
299	41	52.6	131	2	Q7U7T5_SYNPX	Q7u7t5 synchococc	372	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
300	41	52.6	154	2	Q8S29_NITEU	Q8s29 nitrosomona	373	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
301	41	52.6	161	2	Q7VD06_PROMA	Q7vd06 prochloroco	374	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
302	41	52.6	174	2	Q64683_MESAU	Q64683 mesocricetu	375	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
303	41	52.6	182	2	Q94Q15_CHIBR	Q94q15 chinchilla	376	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
304	41	52.6	182	2	Q9B594_9HYST	Q9b594 lagidium vi	377	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
305	41	52.6	192	2	Q94QH4_AULBO	Q94qh4 auliscomys	378	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
306	41	52.6	205	2	Q27672_9NEOP	Q27672 korscheltel	379	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
307	41	52.6	205	2	Q27672_9NEOP	Q2771 sthenopsis a	380	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
308	41	52.6	218	2	Q9Z871_9NEOP	Q9z871 chlamydia p	381	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
309	41	52.6	248	2	Q9ADH0_STRICO	Q9adh0 streptomyce	382	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
310	41	52.6	267	2	Q71P27_9RODE	Q71p27 oxymycterus	383	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
311	41	52.6	267	2	Q71P27_9RODE	Q7m9p2 wolinnella s	384	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
312	41	52.6	284	2	Q93NG1_ARTINI	Q93ng1 artirobacte	385	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
313	41	52.6	306	2	Q8TWE7_METKA	Q8twe7 methanopyru	386	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
314	41	52.6	309	2	Q9SSR0_9ARATH	Q9ssr0 arabidopsis	387	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
315	41	52.6	321	2	Q4S830_TETNG	Q4s830 tetraodon n	388	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
316	41	52.6	327	2	Q8M086_PROWN	Q8m086 prococentru	389	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
317	41	52.6	351	2	Q5D892_SCHJA	Q5des2 schistosoma	390	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
318	41	52.6	356	2	Q8M087_PROWN	Q8m087 prococentru	391	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
319	41	52.6	371	2	Q63H24_BURPS	Q63h24 burkholderi	392	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
320	41	52.6	371	2	Q629U8_BURPA	Q629u8 burkholderi	393	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
321	41	52.6	379	2	Q9T618_LEPCA	Q9t618 lepus capen	394	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
322	41	52.6	379	2	Q9T611_LEPMN	Q9t611 lepus mands	395	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu
323	41	52.6	379	2	Q9T615_LEPSI	Q9t615 lepus sinen	396	40	51.3	972	2	Q5ZMR6_MOUSE	Q5zmr6 mus musculu

397	40	51.3	121	2	Q60GKO_KATPE	Q60gko_katsuwonus	470	40	51.3	173	2	Q85AK8_9SAUR	Q85ak8_liolaemus e
398	40	51.3	125	2	Q9T612_SCOAU	Q9t612_scomber aus	471	40	51.3	173	2	Q85AK7_9SAUR	Q85ak7_liolaemus s
399	40	51.3	127	2	Q79663_SCOAU	Q79663_scomber aus	472	40	51.3	173	2	Q85CH1_9SAUR	Q85ch1_liolaemus p
400	40	51.3	127	2	Q9T611_SCOAU	Q9t611_scomber aus	473	40	51.3	173	2	Q85CH0_9SAUR	Q85ch0_liolaemus p
401	40	51.3	128	2	Q9G280_9HYME	Q9g2n0_calomyrmex	474	40	51.3	173	2	Q85AY8_9SAUR	Q85ay8_liolaemus p
402	40	51.3	128	2	P92842_9HYME	P92842_lysiptalebus	475	40	51.3	173	2	Q85CB4_9SAUR	Q85cb4_liolaemus p
403	40	51.3	128	2	Q7YA05_MICAG	Q7ya05_microtus ag	476	40	51.3	173	2	Q85AD8_9SAUR	Q85ad8_liolaemus s
404	40	51.3	128	2	Q7YA04_MICAG	Q7ya04_microtus ag	477	40	51.3	174	2	Q85DJ4_APOSY	Q85dj4_apodemus sy
405	40	51.3	128	2	Q7Y7D3_MICAG	Q7y7d3_microtus ag	478	40	51.3	174	2	Q85DJ0_APOSY	Q85dj0_apodemus sy
406	40	51.3	128	2	Q7Y7I2_MICAG	Q7y7i2_microtus ag	479	40	51.3	176	1	CYB_IDIPH	Q33537_idionycteri
407	40	51.3	133	2	Q9MGJ7_MICLO	Q9mgj7_microtus lo	480	40	51.3	176	1	CYB_SCICA	Q33537_idionycteri
408	40	51.3	136	2	Q9MGJ6_MICLO	Q9mgj6_microtus lo	481	40	51.3	176	2	Q656A5_ORYSA	Q35805_sciurus car
409	40	51.3	136	2	Q9MGJ8_MICLO	Q9mgj8_microtus lo	482	40	51.3	176	2	Q36131_TAMST	Q656a5_oryza sativ
410	40	51.3	136	2	Q9MGJ9_MICLO	Q9mgj9_microtus lo	483	40	51.3	176	2	Q35955_9SCIU	Q36131_tamias stri
411	40	51.3	136	2	Q9ME21_MICLO	Q9me21_microtus lo	484	40	51.3	176	2	Q34105_CYNGU	Q35955_spermophilu
412	40	51.3	136	2	Q9MDJ6_MICLO	Q9mdj6_microtus lo	485	40	51.3	176	2	Q33725_9SCIU	Q34105_cynomys gun
413	40	51.3	138	2	Q9T614_SCOSE	Q9t614_scomber sco	486	40	51.3	176	2	Q34983_9SCIU	Q33725_amospermop
414	40	51.3	138	1	YMF28_MARPO	P38470_marchantia	487	40	51.3	176	2	Q35913_SPEBP	Q34983_microsciuru
415	40	51.3	139	2	Q79661_RASKA	Q79661_rastrellige	488	40	51.3	176	2	Q34495_TAMQU	Q34495_tamias quad
416	40	51.3	140	2	Q85C83_9CETA	Q85c83_platanista	489	40	51.3	176	2	Q35931_9SCIU	Q35931_sciurus str
417	40	51.3	140	2	Q85P01_MOUSE	Q85p01_mus musculus	490	40	51.3	176	2	Q35824_SPEPR	Q35824_spermophilu
418	40	51.3	140	2	Q85P00_MESAU	Q85p00_mesocricetu	491	40	51.3	176	2	Q33714_9SCIU	Q35824_spermophilu
419	40	51.3	140	2	Q68GV8_9HYST	Q68gv8_hystrix ind	492	40	51.3	176	2	Q34482_TAMMI	Q33714_amospermop
420	40	51.3	140	2	Q9T613_SCOJP	Q9t613_scomber jap	493	40	51.3	176	2	Q35894_SCINI	Q34482_tamias mini
421	40	51.3	140	2	Q79664_SCOAU	Q79664_scomber aus	494	40	51.3	176	2	Q34415_TAMCA	Q35894_sciurus nig
422	40	51.3	140	2	Q79662_SCOJP	Q79662_scomber jap	495	40	51.3	176	2	Q34423_TAMDO	Q34415_tamias cani
423	40	51.3	144	2	Q6PVU8_9HYME	Q6pvu8_trigona hoc	496	40	51.3	176	2	Q34131_CYNLU	Q34423_tamias dors
424	40	51.3	151	2	Q4VDK9_9PERC	Q4vdk9_ichthyys l	497	40	51.3	176	2	Q7GG55_GLASA	Q34131_cynomys lud
425	40	51.3	155	2	Q9T426_PHYDA	Q9t426_phyllotis d	498	40	51.3	177	2	Q85DU2_APOSY	Q7GG55_glaucomya s
426	40	51.3	163	2	Q9G987_VESCR	Q9g987_vespa crabr	499	40	51.3	178	2	Q9B595_CHILA	Q85du2_apodemus sy
427	40	51.3	169	2	Q94Q11_9HYST	Q94q11_octomys mim	500	40	51.3	178	2	Q4VMI3_XROBE	Q9b595_chinchilla
428	40	51.3	170	2	Q8HIW2_BABER	Q8hiw2_barbastella	501	40	51.3	179	2	Q94Q16_CHIBR	Q4vmi3_thalpomys l
429	40	51.3	170	2	Q8HIH3_PEARA	Q8hih3_plecotus au	502	40	51.3	180	2	Q9T2F3_LAGOB	Q4vq16_chinchilla
430	40	51.3	170	2	Q8HIH1_9CHIR	Q8hih1_plecotus te	503	40	51.3	180	2	Q9T2F1_LAGOB	Q9t2f3_lagenorhync
431	40	51.3	170	2	Q8HIH0_9CHIR	Q8hih0_plecotus te	504	40	51.3	180	2	Q9T2F0_LAGOL	Q9t2f1_lagenorhync
432	40	51.3	170	2	Q8HIG9_9CHIR	Q8hig9_plecotus te	505	40	51.3	180	2	Q85DJ9_APOSY	Q9t2f0_lagenorhync
433	40	51.3	170	2	Q8HBV5_9CHIR	Q8hbv5_plecotus te	506	40	51.3	181	2	Q9XNN8_LAGAC	Q85dj9_apodemus sy
434	40	51.3	170	2	Q8HBV2_BABER	Q8hbv2_barbastella	507	40	51.3	181	2	Q9XNN9_LAGAL	Q9xnn8_lagenorhync
435	40	51.3	173	2	Q9T2F4_LAGAC	Q9t2f4_lagenorhync	508	40	51.3	181	2	Q85GL0_9HYST	Q9xnn9_lagenorhync
436	40	51.3	173	2	Q9T2F5_LAGAL	Q9t2f5_lagenorhync	509	40	51.3	181	2	Q3MAZ9_9SAUR	Q85gl0_cavia techu
437	40	51.3	173	2	Q85ID5_9SAUR	Q85id5_liolaemus p	510	40	51.3	181	2	Q3MAY4_9SAUR	Q3maz9_tarentola b
438	40	51.3	173	2	Q85ID4_9SAUR	Q85id4_liolaemus e	511	40	51.3	181	2	Q3MAY2_9SAUR	Q3may4_tarentola b
439	40	51.3	173	2	Q85ID2_9SAUR	Q85id2_liolaemus p	512	40	51.3	181	2	Q3MAY0_9SAUR	Q3may2_tarentola b
440	40	51.3	173	2	Q85IC9_9SAUR	Q85ic9_liolaemus v	513	40	51.3	181	2	Q3MAY1_9SAUR	Q3may0_tarentola b
441	40	51.3	173	2	Q85IC8_9SAUR	Q85ic8_liolaemus s	514	40	51.3	182	2	Q7YFF3_XARMU	Q5may1_tarentola b
442	40	51.3	173	2	Q85IC7_9SAUR	Q85ic7_liolaemus s	515	40	51.3	182	2	Q94Q13_CHILA	Q7yff3_marmosa mur
443	40	51.3	173	2	Q85IC6_9SAUR	Q85ic6_liolaemus s	516	40	51.3	182	2	Q94Q14_CHILA	Q94q13_chinchilla
444	40	51.3	173	2	Q85IC5_9SAUR	Q85ic5_liolaemus s	517	40	51.3	182	2	Q94Q12_CHILA	Q94q12_chinchilla
445	40	51.3	173	2	Q85IC3_9SAUR	Q85ic3_liolaemus k	518	40	51.3	182	2	Q9B592_CHIBR	Q9b592_chinchilla
446	40	51.3	173	2	Q85IC1_9SAUR	Q85ic1_liolaemus p	519	40	51.3	182	2	Q8HH86_CHILA	Q8bh86_chinchilla
447	40	51.3	173	2	Q85IB9_9SAUR	Q85ib9_liolaemus p	520	40	51.3	182	2	Q4VTW0_9SAUR	Q8htw0_gymnodactyl
448	40	51.3	173	2	Q85IB8_9SAUR	Q85ib8_liolaemus e	521	40	51.3	185	2	Q9T2F2_LAGOB	Q9t2f2_lagenorhync
449	40	51.3	173	2	Q85IB7_9SAUR	Q85ib7_liolaemus l	522	40	51.3	185	2	Q68K94_AMMLU	Q68k94_amospermop
450	40	51.3	173	2	Q85IB6_9SAUR	Q85ib6_liolaemus s	523	40	51.3	185	2	Q68K85_AMMLU	Q68k85_amospermop
451	40	51.3	173	2	Q85ID1_9SAUR	Q85id1_liolaemus s	524	40	51.3	185	2	Q68K71_AMMLU	Q68k71_amospermop
452	40	51.3	173	2	Q85CA5_9SAUR	Q85ca5_liolaemus s	525	40	51.3	185	2	Q68KE1_AMMLU	Q68ke1_amospermop
453	40	51.3	173	2	Q85C90_9SAUR	Q85c90_liolaemus s	526	40	51.3	185	2	Q68KA9_AMMLU	Q68ka9_amospermop
454	40	51.3	173	2	Q85B28_9SAUR	Q85b28_liolaemus p	527	40	51.3	185	2	Q68K91_AMMLU	Q68k91_amospermop
455	40	51.3	173	2	Q85APO_9SAUR	Q85apo_liolaemus e	528	40	51.3	185	2	Q9Z2K5_9SAUR	Q9z2k5_uta squamat
456	40	51.3	173	2	Q85AN9_9SAUR	Q85an9_liolaemus b	529	40	51.3	185	2	Q9T510_UTAST	Q9t510_uta stanabu
457	40	51.3	173	2	Q85AN8_9SAUR	Q85an8_liolaemus e	530	40	51.3	185	2	Q9T520_UTAST	Q9t520_uta stanabu
458	40	51.3	173	2	Q85AN7_9SAUR	Q85an7_liolaemus e	531	40	51.3	185	2	Q9Z2K9_9SAUR	Q9z2k9_urosaurus m
459	40	51.3	173	2	Q85AN6_9SAUR	Q85an6_liolaemus k	532	40	51.3	185	2	Q9T509_UTAST	Q9t509_uta stanabu
460	40	51.3	173	2	Q85AN5_9SAUR	Q85an5_liolaemus s	533	40	51.3	185	2	Q9T517_UTAST	Q9t517_uta stanabu
461	40	51.3	173	2	Q85AN4_9SAUR	Q85an4_liolaemus s	534	40	51.3	185	2	Q9Z2K7_9SAUR	Q9z2k7_uta palmeri
462	40	51.3	173	2	Q85AN3_9SAUR	Q85an3_liolaemus s	535	40	51.3	185	2	Q9T515_UTAST	Q9t515_uta stanabu
463	40	51.3	173	2	Q85AN2_9SAUR	Q85an2_liolaemus s	536	40	51.3	185	2	Q9Z2K3_UTAST	Q9z2k3_uta stanabu
464	40	51.3	173	2	Q85AN1_9SAUR	Q85an1_liolaemus s	537	40	51.3	185	2	Q9T519_UTAST	Q9t519_uta stanabu
465	40	51.3	173	2	Q85AN0_9SAUR	Q85an0_liolaemus k	538	40	51.3	185	2	Q9T518_UTAST	Q9t518_uta stanabu
466	40	51.3	173	2	Q85C14_9SAUR	Q85c14_liolaemus p	539	40	51.3	185	2	Q9T516_UTAST	Q9t516_uta stanabu
467	40	51.3	173	2	Q85C19_9SAUR	Q85c19_liolaemus e	540	40	51.3	185	2	Q9T514_UTAST	Q9t514_uta stanabu
468	40	51.3	173	2	Q85C18_9SAUR	Q85c18_liolaemus s	541	40	51.3	185	2	Q9T512_UTAST	Q9t512_uta stanabu
469	40	51.3	173	2	Q85C17_9SAUR	Q85c17_liolaemus s	542	40	51.3	185	2	Q9Z2K8_9SAUR	Q9z2k8_petrosaurus

543	40	51.3	185	2	Q9T508_UTAST	Q9T508_uta stanebu	616	40	51.3	195	2	Q85QY5_LEPCA	Q85QY5 lepus capen
544	40	51.3	185	2	Q92ZK3_9SAUR	Q92zk2_uta stefneg	617	40	51.3	195	2	Q85QY3_LEPCA	Q85QY3 lepus capen
545	40	51.3	185	2	Q9T511_UTAST	Q9T513_uta stanebu	618	40	51.3	195	2	Q85QY2_LEPCA	Q85QY2 lepus capen
546	40	51.3	185	2	Q92ZK6_9SAUR	Q92zk6_uta stellat	619	40	51.3	195	2	Q85QY1_LEPSX	Q85QY1 lepus saxat
547	40	51.3	185	2	Q92ZK4_9SAUR	Q92zk4_uta antiqua	620	40	51.3	195	2	Q85QY0_LEPSX	Q85QY0 lepus saxat
548	40	51.3	185	2	Q9T511_UTAST	Q9T511_uta stanebu	621	40	51.3	195	2	Q85BA6_LEPGR	Q85BA6 lepus grana
549	40	51.3	185	2	Q9T3A3_UTAST	Q9T3a3_uta stanebu	622	40	51.3	195	2	Q85B32_LEPGR	Q85B32 lepus grana
550	40	51.3	185	2	Q9T3A2_UTAST	Q9T3a2_uta stanebu	623	40	51.3	195	2	Q85B31_LEPEU	Q85B31 lepus europ
551	40	51.3	185	2	Q4VTV4_9SAUR	Q4vtv4_gymnodactyl	624	40	51.3	195	2	Q36161_THUTH	Q36161 thunnus thy
552	40	51.3	186	2	Q9XNN7_LAGOL	Q9xnn7_lagenorhynch	625	40	51.3	196	2	Q33879_9SCOM	Q33879 auxis thaza
553	40	51.3	186	2	Q85QZ9_LEPEU	Q85qz9 lepus europ	626	40	51.3	196	2	Q34405_9SCOM	Q34405 euthynnus a
554	40	51.3	186	2	Q5VC40_9RODE	Q5vc40_peromyscus	627	40	51.3	196	2	Q34403_9SCOM	Q34403 gymnodactyl
555	40	51.3	186	2	Q5VBUI_9RODE	Q5vbui_peromyscus	628	40	51.3	197	2	Q4VTX0_9SAUR	Q4vtx0_gymnodactyl
556	40	51.3	186	2	Q5VBS9_9RODE	Q5vbs9_peromyscus	629	40	51.3	198	2	Q85LL0_9SCOM	Q85LL0 auxis thaza
557	40	51.3	186	2	Q5VBQ3_9RODE	Q5vbq1_peromyscus	630	40	51.3	198	2	Q85LK7_9SCOM	Q85LK7 euthynnus a
558	40	51.3	186	2	Q8HBP8_CHIBR	Q8hbp8_chinchilla	631	40	51.3	198	2	Q85LK5_KATPE	Q85LK5 katsuonon
559	40	51.3	186	2	Q5VCA6_9RODE	Q5vca6_peromyscus	632	40	51.3	198	2	Q85LK3_SAROR	Q85LK3 sarda orien
560	40	51.3	186	2	Q5VBM2_9RODE	Q5vbm2_peromyscus	633	40	51.3	198	2	Q85LK1_THUAA	Q85LK1 thunnus alia
561	40	51.3	186	2	Q5VBP8_9RODE	Q5vbp8_peromyscus	634	40	51.3	198	2	Q85LJ8_THUAL	Q85LJ8 thunnus alb
562	40	51.3	186	2	Q5VBS6_9RODE	Q5vbs6_peromyscus	635	40	51.3	198	2	Q85CG0_AUXRO	Q85CG0 auxis roche
563	40	51.3	186	2	Q5VBL7_9RODE	Q5vbl7_peromyscus	636	40	51.3	198	2	Q85CF9_KATPE	Q85CF9 katsuonon
564	40	51.3	186	2	Q5VBS3_9RODE	Q5vbs3_peromyscus	637	40	51.3	198	2	Q4VTX6_9SAUR	Q4vtx6_gymnodactyl
565	40	51.3	186	2	Q5VBT9_9RODE	Q5vbt9_peromyscus	638	40	51.3	198	2	Q4VTV6_9SAUR	Q4vtv6_gymnodactyl
566	40	51.3	186	2	Q5VBL0_9RODE	Q5vbl0_peromyscus	639	40	51.3	199	1	CXB_SARCH	P34864 sarda chilii
567	40	51.3	186	2	Q5VBM9_9RODE	Q5vbm9_peromyscus	640	40	51.3	199	2	Q7YFF2_MARMU	Q7yff2 marmosa mur
568	40	51.3	186	2	Q5VBS7_9RODE	Q5vbs7_peromyscus	641	40	51.3	199	2	Q35975_THUAA	Q35975 thunnus alia
569	40	51.3	187	2	Q8SQY7_LEPTI	Q8sqy7 lepus timid	642	40	51.3	199	2	Q35867_SCOJP	Q35867 scomber jap
570	40	51.3	187	2	Q6TQM3_9SAUR	Q6tqm3_liolaemus g	643	40	51.3	199	2	Q34836_KATPE	Q34836 katsuonon
571	40	51.3	187	2	Q4VTY4_9SAUR	Q4vtx4_gymnodactyl	644	40	51.3	199	2	Q34590_9SCOM	Q34590 gastereochis
572	40	51.3	187	2	Q4VTX9_9SAUR	Q4vtx9_gymnodactyl	645	40	51.3	200	2	Q92ZG3_9RODE	Q92zg3 phyllotis m
573	40	51.3	187	2	Q8SQY6_LEPTI	Q8sqy6 lepus timid	646	40	51.3	200	2	Q85A48_CLEGL	Q85a48 clethrionom
574	40	51.3	190	2	Q8MGJ4_9SAUR	Q8mgj4_darevskia v	647	40	51.3	200	2	Q52TK6_CLEGA	Q52tk6 clethrionom
575	40	51.3	190	2	Q4VTV1_9SAUR	Q4vtv1_gymnodactyl	648	40	51.3	200	2	Q52TJ4_CLEGA	Q52tj4 clethrionom
576	40	51.3	191	2	Q85QZ8_LEPEU	Q85qz8 lepus europ	649	40	51.3	200	2	Q52TJ2_CLEGA	Q52tj2 clethrionom
577	40	51.3	191	2	Q85B70_LEPGR	Q85b70 lepus grana	650	40	51.3	200	2	Q52TJ0_CLEGA	Q52tj0 clethrionom
578	40	51.3	192	2	Q33689_9CETA	Q33689_australopho	651	40	51.3	200	2	Q52T19_CLEGA	Q52t19 clethrionom
579	40	51.3	192	2	Q35164_NEOPH	Q35164_neophocaena	652	40	51.3	200	2	Q52T18_CLEGA	Q52t18 clethrionom
580	40	51.3	192	2	Q35628_PHOSS	Q35628_phocaena si	653	40	51.3	200	2	Q52T13_CLEGA	Q52t13 clethrionom
581	40	51.3	192	2	Q35603_PHOPH	Q35603_phocoenoid	654	40	51.3	200	2	Q52TH5_CLEGA	Q52th5 clethrionom
582	40	51.3	192	2	Q85R09_LEPGR	Q85r09 lepus grana	655	40	51.3	200	2	Q52TJ5_CLEGA	Q52tj5 clethrionom
583	40	51.3	192	2	Q85R09_LEPGR	Q85qz1 lepus corai	656	40	51.3	200	2	Q52T11_CLEGA	Q52t11 clethrionom
584	40	51.3	192	2	Q37055_9CETA	Q37055_phocoenoid	657	40	51.3	200	2	Q52TH6_CLEGA	Q52th6 clethrionom
585	40	51.3	192	2	Q9MPW6_CHILA	Q9mpw6_chinchilla	658	40	51.3	200	2	Q52TM9_CLEGA	Q52tm9 clethrionom
586	40	51.3	192	2	Q6X4W7_9HXT	Q6x4w7_lagidium vi	659	40	51.3	200	2	Q52TK9_CLEGA	Q52tk9 clethrionom
587	40	51.3	193	2	Q85R10_LEPGR	Q85r10 lepus grana	660	40	51.3	200	2	Q35976_THUAL	Q35976 thunnus alb
588	40	51.3	193	2	Q85QZ7_LEPEU	Q85qz7 lepus europ	661	40	51.3	201	2	Q85IA4_9SAUR	Q85ia4 liolaemus e
589	40	51.3	193	2	Q85QZ4_LEPEU	Q85qz4 lepus europ	662	40	51.3	201	2	Q9G4Q6_9RODE	Q9g4q6 peromyscus
590	40	51.3	193	2	Q85QZ2_LEPCS	Q85qz2 lepus castr	663	40	51.3	201	2	Q33861_9SCOM	Q33861 acanthocybi
591	40	51.3	193	2	Q85QY9_LEPTI	Q85qy9 lepus timid	664	40	51.3	201	2	Q85I76_9SAUR	Q85i76 liolaemus s
592	40	51.3	193	2	Q6X4W6_9HXT	Q6x4w6_lagidium vi	665	40	51.3	202	2	Q36093_THUOB	Q36093 thunnus obe
593	40	51.3	194	2	Q85R13_LEPGR	Q85r13 lepus grana	666	40	51.3	202	2	Q6TQV6_9SAUR	Q6tqv6 liolaemus g
594	40	51.3	194	2	Q85R05_LEPGR	Q85r05 lepus grana	667	40	51.3	202	2	Q35922_SCOSC	Q35922 scomber sco
595	40	51.3	194	2	Q85R04_LEPGR	Q85r04 lepus grana	668	40	51.3	203	2	Q85I99_9SAUR	Q85i99 liolaemus e
596	40	51.3	194	2	Q85R03_LEPGR	Q85r03 lepus grana	669	40	51.3	203	2	Q27639_9NROP	Q27639 epimartyria
597	40	51.3	194	2	Q5IXI6_9SAUR	Q5ixi6_thecadactyl	670	40	51.3	204	2	Q4VTY3_9SAUR	Q4vtcy3 gymnodactyl
598	40	51.3	194	2	Q5IXG9_9SAUR	Q5ixg9_thecadactyl	671	40	51.3	205	2	Q9MG44_9SCOM	Q9mg44 thunnus atl
599	40	51.3	194	2	Q5IXG8_9SAUR	Q5ixg8_thecadactyl	672	40	51.3	205	2	Q85L11_AUXRO	Q85l11 auxis roche
600	40	51.3	194	2	Q5IXG2_9SAUR	Q5ixg2_thecadactyl	673	40	51.3	205	2	Q85L11_9SCOM	Q85l11 auxis roche
601	40	51.3	194	2	Q5IXG1_9SAUR	Q5ixg1_thecadactyl	674	40	51.3	205	2	Q85LK6_9SCOM	Q85lk6 euthynnus a
602	40	51.3	194	2	Q4VTW2_9SAUR	Q4vtw2_gymnodactyl	675	40	51.3	205	2	Q85LK4_KATPE	Q85lk4 katsuonon
603	40	51.3	195	2	Q85R14_LEPGR	Q85r14 lepus grana	676	40	51.3	205	2	Q85LK2_SAROR	Q85lk2 sarda orien
604	40	51.3	195	2	Q85R12_LEPGR	Q85r12 lepus grana	677	40	51.3	205	2	Q85LJ7_THUAL	Q85lj7 thunnus alb
605	40	51.3	195	2	Q85R11_LEPGR	Q85r11 lepus grana	678	40	51.3	205	2	Q85LJ4_9SCOM	Q85lj4 thunnus mac
606	40	51.3	195	2	Q85R07_LEPGR	Q85r07 lepus grana	679	40	51.3	205	2	Q85LJ3_THUOB	Q85lj3 thunnus obe
607	40	51.3	195	2	Q85R06_LEPGR	Q85r06 lepus grana	680	40	51.3	205	2	Q85LJ0_THUOB	Q85lj0 thunnus obe
608	40	51.3	195	2	Q85R02_LEPGR	Q85r02 lepus grana	681	40	51.3	205	2	Q85L18_THUTO	Q85l18 thunnus thy
609	40	51.3	195	2	Q85R01_LEPGR	Q85r01 lepus grana	682	40	51.3	205	2	Q85L16_THUTH	Q85l16 thunnus ton
610	40	51.3	195	2	Q85R00_LEPEU	Q85r00 lepus europ	683	40	51.3	205	2	Q9MG43_9SCOM	Q9mg43 thunnus ton
611	40	51.3	195	2	Q85QZ6_LEPEU	Q85qz6 lepus europ	684	40	51.3	206	2	Q36069_9SCOM	Q36069 thunnus mac
612	40	51.3	195	2	Q85QZ5_LEPEU	Q85qz5 lepus europ	685	40	51.3	206	2	Q6TQY1_9SAUR	Q6tqy1 liolaemus s
613	40	51.3	195	2	Q85QZ3_LEPCS	Q85qz3 lepus castr	686	40	51.3	206	2	Q6TQV5_9SAUR	Q6tqv5 liolaemus g
614	40	51.3	195	2	Q85QZ0_LEPTI	Q85qz0 lepus timid	687	40	51.3	207	2	Q6X4W9_9HYST	Q6x4w9 lagidium pe
615	40	51.3	195	2	Q85QY8_LEPTI	Q85qy8 lepus timid	688	40	51.3	207	2	Q6X4W8_9HYST	Q6x4w8 lagidium vi

689	40	51.3	207	2	Q6X4W5	lagidium vi	Q6x4w5	762	40	51.3	221	2	Q69FU7	9RODE	Q69fu7	arborimus 1
690	40	51.3	207	2	Q85IB0	liolaemus e	Q85ib0	763	40	51.3	221	2	Q69FS8	9RODE	Q69fs8	arborimus 1
691	40	51.3	207	2	Q85IA7	liolaemus e	Q85ia7	764	40	51.3	221	2	Q69FR6	9RODE	Q69fr6	arborimus a
692	40	51.3	207	2	Q8TR01	liolaemus d	Q8tr01	765	40	51.3	221	2	Q69FR3	9PHEIN	Q69fr3	phenacomys
693	40	51.3	208	2	Q21639	darevskia d	Q21639	766	40	51.3	221	2	Q69FR2	9PHEIN	Q69fr2	phenacomys
694	40	51.3	208	2	Q21632	darevskia a	Q21632	767	40	51.3	221	2	Q69FR1	9RODE	Q69fr1	arborimus p
695	40	51.3	209	2	Q58J79	lepus grana	Q58j79	768	40	51.3	221	2	Q69FW5	9RODE	Q69fw5	arborimus 1
696	40	51.3	209	2	Q58J78	lepus grana	Q58j78	769	40	51.3	221	2	Q69FV1	9RODE	Q69fv1	arborimus 1
697	40	51.3	209	2	Q58J77	lepus europ	Q58j77	770	40	51.3	221	2	Q69FU2	9RODE	Q69fu2	arborimus 1
698	40	51.3	209	2	Q58J76	lepus grana	Q58j76	771	40	51.3	221	2	Q69FU1	9RODE	Q69fu1	arborimus 1
699	40	51.3	209	2	Q58J75	lepus castr	Q58j75	772	40	51.3	221	2	Q69FQ9	9RODE	Q69fq9	arborimus p
700	40	51.3	209	2	Q52N14	mabuya macu	Q52n14	773	40	51.3	221	2	Q69FV9	9RODE	Q69fv9	arborimus 1
701	40	51.3	210	2	Q4VTW6	gymnodactyl	Q4vtw6	774	40	51.3	221	2	Q69FR8	9RODE	Q69fr8	arborimus a
702	40	51.3	211	2	Q85LK0	thunus ala	Q85lk0	775	40	51.3	221	2	Q957I9	LACAG	Q957i9	lacerta agi
703	40	51.3	211	2	Q85LK9	thunus alb	Q85lk9	776	40	51.3	221	2	Q957I8	LACBL	Q957i8	lacerta bil
704	40	51.3	211	2	Q85LK6	thunus atl	Q85lk6	777	40	51.3	221	2	Q8M2D4	9SAUR	Q8m2d4	pachydactyl
705	40	51.3	211	2	Q85LK5	thunus mac	Q85lk5	778	40	51.3	221	2	Q8M2E4	9SAUR	Q8m2e4	pachydactyl
706	40	51.3	211	2	Q85LK2	thunus obe	Q85lk2	779	40	51.3	221	2	Q8M2E1	9SAUR	Q8m2e1	pachydactyl
707	40	51.3	211	2	Q85LK1	thunus obe	Q85lk1	780	40	51.3	221	2	Q952G5	9SAUR	Q952g5	lacerta sch
708	40	51.3	211	2	Q85L19	thunus thy	Q85l19	781	40	51.3	221	2	Q952G3	9SAUR	Q952g3	lacerta sch
709	40	51.3	211	2	Q85L17	thunus ton	Q85l17	782	40	51.3	221	2	Q8M2D2	9SAUR	Q8m2d2	pachydactyl
710	40	51.3	211	2	Q85L15	thunus ton	Q85l15	783	40	51.3	221	2	Q952G4	9SAUR	Q952g4	lacerta sch
711	40	51.3	212	2	Q85L12	auxis roche	Q85l12	784	40	51.3	221	2	Q8M2E2	9SAUR	Q8m2e2	pachydactyl
712	40	51.3	212	2	Q6TRI0	liolaemus d	Q6tri0	785	40	51.3	221	2	Q6TQV5	9SAUR	Q6tgV5	liolaemus 1
713	40	51.3	212	2	Q52N09	mabuya macu	Q52n09	786	40	51.3	221	2	Q94PE2	9SAUR	Q94pe2	lacerta sch
714	40	51.3	213	2	Q69FV4	arborimus 1	Q69fv4	787	40	51.3	221	2	Q94PE1	9SAUR	Q94pe1	lacerta sch
715	40	51.3	213	2	Q9MGJ3	darevskia p	Q9mgj3	788	40	51.3	221	2	Q94PE0	9SAUR	Q94pe0	lacerta sch
716	40	51.3	213	2	Q7YGM2	kenopia str	Q7ygm2	789	40	51.3	221	2	Q94PK3	9SAUR	Q94pk3	lacerta sch
717	40	51.3	213	2	Q6TR40	liolaemus d	Q6tr40	790	40	51.3	223	2	Q6TRO8	9SAUR	Q6trO8	liolaemus d
718	40	51.3	213	2	Q6TQV0	liolaemus s	Q6tqv0	791	40	51.3	223	2	Q6TQU8	9SAUR	Q6tqu8	liolaemus u
719	40	51.3	213	2	Q6TQV3	liolaemus g	Q6tqv3	792	40	51.3	223	2	Q4VTW7	9SAUR	Q4vtw7	gymnodactyl
720	40	51.3	213	2	Q52N10	mabuya macu	Q52n10	793	40	51.3	223	2	Q4VTM4	9SAUR	Q4vtM4	gymnodactyl
721	40	51.3	214	2	Q85TD9	cavia tschu	Q85td9	794	40	51.3	224	2	Q9XMJ9	BRAID	Q9xmJ9	brachylagus
722	40	51.3	215	2	Q34872	oryctolagus	Q34872	795	40	51.3	224	2	Q9THE7	SYLFL	Q9the7	syvilagus
723	40	51.3	215	2	Q85TD3	galia muete	Q85td3	796	40	51.3	224	2	Q9XMK1	LEPCL	Q9xmK1	lepus calif
724	40	51.3	215	2	Q52N11	mabuya macu	Q52n11	797	40	51.3	224	2	Q9XMK2	LEPCA	Q9xmK2	lepus capen
725	40	51.3	215	2	Q52N03	mabuya macu	Q52n03	798	40	51.3	224	2	Q9XMK5	SYLAQ	Q9xmK5	syvilagus
726	40	51.3	216	2	Q9TAG3	una paraphy	Q9tag3	799	40	51.3	224	2	Q9XMK0	BUNMO	Q9xmK0	bunolagus m
727	40	51.3	216	2	Q9TAGV	una exsul.	Q9tag4	800	40	51.3	224	2	Q9THE6	LEPAM	Q9the6	lepus ameri
728	40	51.3	217	2	Q4VTY0	gymnodactyl	Q4vty0	801	40	51.3	224	2	Q9XMK4	ROMDI	Q9xmK4	romerolagus
729	40	51.3	217	2	Q9TAF7	holbrookia	Q9taf7	802	40	51.3	225	2	Q9TGR9	9SAUR	Q9tgr9	lepus timid
730	40	51.3	217	2	Q9TAF3	holbrookia	Q9taf3	803	40	51.3	225	2	Q85I90	9SAUR	Q85i90	liolaemus e
731	40	51.3	217	2	Q9TAG7	cophesaurus	Q9tag7	804	40	51.3	225	2	Q7YAZ3	9SAUR	Q7yaz3	cordyllosaur
732	40	51.3	217	2	Q9TAG6	cophesaurus	Q9tag6	805	40	51.3	225	2	Q52N21	9SAUR	Q52n21	mabuya macu
733	40	51.3	217	2	Q9TAG2	holbrookia	Q9tag2	806	40	51.3	225	2	Q52N20	9SAUR	Q52n20	mabuya macu
734	40	51.3	217	2	Q9TAF4	holbrookia	Q9taf4	807	40	51.3	225	2	Q52N05	9SAUR	Q52n05	mabuya macu
735	40	51.3	217	2	Q9TAF0	holma	Q9taf0	808	40	51.3	225	2	Q4VTV9	9SAUR	Q4vtv9	gymnodactyl
736	40	51.3	217	2	Q9TAH3	urosaurs o	Q9tah3	809	40	51.3	226	2	Q71FD4	BARBR	Q71fd4	barbastella
737	40	51.3	217	2	Q9TAH4	UTAST	Q9tah4	810	40	51.3	226	2	Q71FC6	BARBR	Q71fc6	barbastella
738	40	51.3	217	2	Q9TAF6	holma	Q9taf6	811	40	51.3	226	2	Q71FC5	CHIR	Q71fc5	barbastella
739	40	51.3	217	2	Q9TAG8	caldr	Q9tag8	812	40	51.3	226	2	Q71FB1	PLEAS	Q71fb1	plecotus au
740	40	51.3	217	2	Q9TAF8	holma	Q9taf8	813	40	51.3	226	2	Q71FA8	PLEAS	Q71fa8	plecotus au
741	40	51.3	217	2	Q9TAG9	calldr	Q9tag9	814	40	51.3	226	2	Q71FA4	PLEAS	Q71fa4	plecotus au
742	40	51.3	217	2	Q9TAF9	holma	Q9taf9	815	40	51.3	226	2	Q71FA3	PLEAS	Q71fa3	plecotus au
743	40	51.3	217	2	Q9TAF5	holma	Q9taf5	816	40	51.3	226	2	Q71FA2	PLEAS	Q71fa2	plecotus au
744	40	51.3	217	2	Q9TAG1	9SAUR	Q9tag1	817	40	51.3	226	2	Q71PA0	CHIR	Q71fa0	plecotus cf
745	40	51.3	217	2	Q9TAG5	9SAUR	Q9tag5	818	40	51.3	226	2	Q71F99	CHIR	Q71f99	plecotus cf
746	40	51.3	217	2	Q6TQV0	9SAUR	Q6tqv0	819	40	51.3	226	2	Q71F97	CHIR	Q71f97	plecotus ko
747	40	51.3	217	2	Q9T454	UMASC	Q9tq454	820	40	51.3	226	2	Q71F96	PLEAS	Q71f96	plecotus au
748	40	51.3	217	2	Q9T4Q0	CALDR	Q9t4q0	821	40	51.3	226	2	Q71F93	CHIR	Q71f93	plecotus sp
749	40	51.3	217	2	Q9T3V7	UMANO	Q9t3v7	822	40	51.3	226	2	Q71F92	CHIR	Q71f92	plecotus sp
750	40	51.3	218	2	Q6TR09	9SAUR	Q6tr09	823	40	51.3	226	2	Q71F89	CHIR	Q71f89	plecotus ba
751	40	51.3	218	2	Q6TQV4	9SAUR	Q6tqv4	824	40	51.3	226	2	Q71F88	CHIR	Q71f88	plecotus ba
752	40	51.3	218	2	Q4VTV8	9SAUR	Q4vtv8	825	40	51.3	226	2	Q71F80	CHIR	Q71f80	plecotus te
753	40	51.3	219	2	Q6KCT8	MICDE	Q6kct8	826	40	51.3	226	2	Q71F79	CHIR	Q71f79	plecotus te
754	40	51.3	219	2	Q9MM61	9SAUR	Q9mm61	827	40	51.3	226	2	Q6W818	9CHIR	Q6w818	plecotus sp
755	40	51.3	219	2	Q6TOX8	9SAUR	Q6tox8	828	40	51.3	226	2	Q71FA5	CHIR	Q71fa5	plecotus sp
756	40	51.3	220	2	Q9RBY2	MARMO	Q9rby2	829	40	51.3	226	2	Q71F91	PLEAS	Q71f91	plecotus au
757	40	51.3	220	2	Q9MIB1	9SAUR	Q9mib1	830	40	51.3	226	2	Q71FD2	BARBR	Q71fd2	barbastella
758	40	51.3	220	2	Q85N08	9SAUR	Q85n08	831	40	51.3	226	2	Q71FC7	BARBR	Q71fc7	barbastella
759	40	51.3	220	2	Q85H40	pachydactyl	Q85h40	832	40	51.3	226	2	Q71FA6	BARBR	Q71fa6	plecotus au
760	40	51.3	220	2	Q6TR56	9SAUR	Q6tr56	833	40	51.3	226	2	Q71F90	PLEAS	Q71f90	plecotus au
761	40	51.3	220	2	Q4VTV5	gymnodactyl	Q4vtv5	834	40	51.3	226	2	Q85TE0	CAVPO	Q85te0	cavia porce

981 40 51.3 238 2 Q9T6I6 LEPOI
982 40 51.3 238 2 Q4JFM0 SPERMOPHILU
983 40 51.3 238 2 Q4JFM1 SPESP
984 40 51.3 238 2 Q03375 PARAXERUS C
985 40 51.3 238 2 Q6XNZ7 PACHYDACTYL
986 40 51.3 238 2 Q6XNZ6 RHOTROPUS
987 40 51.3 238 2 Q6XNZ3 PHELSUMA MU
988 40 51.3 238 2 Q6XNZ2 PHELSUMA AS
989 40 51.3 238 2 Q6XNY9 PHELSUMA MA
990 40 51.3 238 2 Q6XNY7 PHELSUMA SE
991 40 51.3 238 2 Q6XNY6 PHELSUMA LI
992 40 51.3 238 2 Q6XNY4 PHELSUMA ST
993 40 51.3 238 2 Q6XNY3 PHELSUMA QU
994 40 51.3 238 2 Q6XNY2 PHELSUMA AB
995 40 51.3 238 2 Q6XNY1 PHELSUMA AB
996 40 51.3 238 2 Q6XNY0 PHELSUMA AB
997 40 51.3 238 2 Q6XNS9 PHELSUMA GU
998 40 51.3 238 2 Q6UKV6 PHELSUMA BU
999 40 51.3 238 2 Q6TR69 LIOLAEMUS D
1000 40 51.3 238 2 Q6TR63 LIOLAEMUS D

ALIGNMENTS

RESULT 1
ID CK6A2_BOVIN STANDARD; PRT; 97 AA.
AC P07471;
DC 01-APR-1988 (Rel. 07, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Cytochrome c oxidase polypeptide Via-heart, mitochondrial precursor
DE (EC 1.9.3.1) (COXVIAH) (Polypeptide VIB).
GN Name=COX6A2; Synonyms=COX6A;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=91274363; PubMed=1647214; DOI=10.1016/0167-4781(91)90022-E;
RA Smith E.O.; Bement D.M.; Grossman L.I.; Lomax M.I.;
RT "The cDNA for the heart/muscle isoform of bovine cytochrome c oxidase
RT subunit V1a encodes a presequence.";
RL Biochim. Biophys. Acta 1089:266-268(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=93326637; PubMed=7687470; DOI=10.1016/0167-4781(93)90092-R;
RA Smith E.O.; Lomax M.I.;
RT "Structural organization of the bovine gene for the heart/muscle
RT isoform of cytochrome c oxidase subunit V1a.";
RL Biochim. Biophys. Acta 1174:63-71(1993).
RN [3]
RP PROTEIN SEQUENCE OF 13-96.
RC TISSUE=Heart;
RX MEDLINE=86000134; PubMed=2994692;
RA Meinecke L.; Buse G.;
RT "Studies on cytochrome c oxidase, XII. Isolation and primary structure
RT of polypeptide V1b from bovine heart.";
RL Biol. Chem. Hoppe-Seyler 366:687-694(1985).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=96216288; PubMed=8638158;
RA Tsukihara T.; Aoyama H.; Yamashita E.; Tomizaki T.; Yamaguchi H.;
RA Shinzawa-Itou K.; Nakashima R.; Yano R.; Yoshikawa S.;
RT "The whole structure of the 13-subunit oxidized cytochrome c oxidase
RT at 2.8 A.";
RL Science 272:1136-1144(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RC TISSUE=Heart;
RX MEDLINE=99190827; PubMed=10089392;
RA Tomizaki T.; Yamashita E.; Yamaguchi H.; Aoyama H.; Tsukihara T.;
RA Shinzawa-Itou K.; Nakashima R.; Yano R.; Yoshikawa S.;
RT "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A
RT resolution.";
RL Acta Crystallogr. D 55:31-45(1999).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Heart;
RX MEDLINE=20235645; PubMed=10771420; DOI=10.1107/S0907444900002213;
RA Fei M.J.; Yamashita E.; Inoue N.; Yao M.; Yamaguchi H.; Tsukihara T.;
RA Shinzawa-Itou K.; Nakashima R.; Yoshikawa S.;
RT "X-ray structure of azide-bound fully oxidized cytochrome c oxidase
RT from bovine heart at 2.9 A resolution.";
RL Acta Crystallogr. D 56:529-535(2000).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- TISSUE SPECIFICITY: Heart/muscle specific isoform.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase Via family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; S64127; AAB27605.1; -; Genomic_DNA.
CC EMBL; X56857; CAA40183.1; -; mRNA.
CC PIR; S35702; OGB06.
CC PDB; 1OCC; X-ray; G/T=13-96.
CC PDB; 1OCO; X-ray; G/T=13-96.
CC PDB; 1OCR; X-ray; G/T=13-96.
CC PDB; 1OCZ; X-ray; G/T=13-96.
CC PDB; 1V54; X-ray; G/T=13-97.
CC PDB; 1V55; X-ray; G/T=13-97.
CC PDB; 2OCC; X-ray; G/T=13-96.
CC InterPro; IPR001349; COX6A.
CC PANTHER; PTHR11504; COX6A; 1.
CC Pfam; PF02046; COX6A; 1.
CC PIRSF; PIRSF00277; COX6A1; 1.
CC PRODOM; PD006036; COX6A; 1.
CC PROSITE; PS01329; COX6A; 1.
CC 3D-structure; Direct protein sequencing; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Transit peptide; Transmembrane.
FT TRANSIT 1 12
FT CHAIN 13 97
FT Cytochrome c oxidase polypeptide Via-
FT heart.
FT TOPO DOM 13 24
FT Mitochondrial matrix.
FT TRANSMEM 25 49
FT TOPO DOM 50 97
FT Mitochondrial intermembrane.
SQ SEQUENCE 97 AA; 10800 MW; 59B408A1B9B6AF4F CRC64;

Query Match 60.3%; Score 47; DB 1; Length 97;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGG 10
||| ||| |||
DB 43 CTLNSWLHSG 52

RESULT 2
ID Q9GAB0_9BIVA PRELIMINARY; PRT; 128 AA.
AC Q9GAB0;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)

DE Cytochrome b (Fragment).
GN Name=cytb;
OS Pitar morrhuanus (false quahog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Pitar.
OX NCBI_TaxID=115818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
CC EMBL; AF205082; AAG32644.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_NTER; 1.
DR ELECTRON_TRANSPORT; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 128 128
FT SEQUENCE 128 AA; 13986 MW; EA78F42E59583BEE CRC64;
SQ
Query Match 60.3%; Score 47; DB 2; Length 128;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 TLREWLHGFC 12
: |||||
Db 18 SLVEWLGGFC 28

RESULT 3
Q9GAB2_MERCA
ID Q9GAB2_MERCA PRELIMINARY; PRT; 128 AA.
AC Q9GAB2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Mercenaria campechiensis (Southern quahog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Mercenaria.
OX NCBI_TaxID=115816;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).

CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205080; AAG32642.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR ELECTRON_TRANSPORT; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 128 128
FT SEQUENCE 128 AA; 14058 MW; 2B3C1546BBC9B8F3 CRC64;
SQ
Query Match 59.0%; Score 46; DB 2; Length 128;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 LREWLHGFC 12
: |||||
Db 19 LVEWLGGFC 28

RESULT 4
Q9GAB1_9BIVA
ID Q9GAB1_9BIVA PRELIMINARY; PRT; 128 AA.
AC Q9GAB1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Calyptogenia sp. 'Kodiak Seep'.
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Gloosoidae; Vesicomysidae; Calyptogenia.
OX NCBI_TaxID=118660;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dahlgren T.G., Weinberg J.R., Halanynch K.M.;
RT "Phylogeography of the ocean quahog (Arctica islandica): influences of
RT paleoclimate on genetic diversity and species range.";
RL Mar. Biol. 137:487-495(2000).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF205081; AAG32643.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR ELECTRON_TRANSPORT; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 128 128
FT SEQUENCE 128 AA; 14058 MW; 2B3C1546BBC9B8F3 CRC64;
SQ

AS Ashbya gossypii (Yeast) (*Eremothecium gossypii*).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

DR InterPro; IPR011498; Kelch_2.
DR Pfam; PF07646; Kelch_2; 2.
DR SMART; SM00387; HATPase_c; 1.
KW Hypothetical protein.
SQ SEQUENCE 2558 AA; 295828 MW; 6B00A5209A3489AB CRC64;

Query Match 59.0%; Score 46; DB 2; Length 2558;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
   |||::|||
Db 1282 CTLQKWLFGAF 1292

RESULT 10
Q7PDT6 PLAYO
ID Q7PDT6_PLAYO PRELIMINARY; PRT; 2584 AA.
AC Q7PDT6;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Erythrocyte membrane protein PFEMP3.
GN Name=PY01412;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angioli S.V., Suh B.B., Kooij T.W., Perte M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin H.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519 (2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABL01000372; EAA20743.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR Pfam; PF02518; HATPase_c; 1.
DR ORFNames=PF14_0649;
SQ SEQUENCE 2584 AA; 297199 MW; D5E87355F355007E CRC64;

Query Match 59.0%; Score 46; DB 2; Length 2584;
Best Local Similarity 63.6%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
   |||::|||
Db 1283 CTLQKWLFGAF 1293

RESULT 11
Q8TJEO METAC
ID Q8TJEO_METAC PRELIMINARY; PRT; 126 AA.
AC Q8TJEO;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=MA3845;
OS Methanosarcina acetivorans.

```

```
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C2A / ATCC 35395 / DSM 2834; DOI=10.1101/gr.223902;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Scange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Metcalf W.W., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AB011095; AA007196.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 126 AA; 14491 MW; 30FC5B8FDEA11926 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 126;
Best Local Similarity 70.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTLREWLHGG 10
|||:|:|
Db 70 CTIREYLHDG 79

RESULT 12
Q9ZNB4_PORGI PRELIMINARY; PRT; 211 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DB 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Insertion sequence IS1126-like gene.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RX MEDLINE=95143166; PubMed=988746; DOI=10.1074/jbc.274.8.5012;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products.";
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL; AB019363; BAA34342.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1-
DR PIR; P01609; Transposase_11; 1-
SQ SEQUENCE 211 AA; 24444 MW; 77040F65D76D1175 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLHGGFC 12
::|:|:|
Db 167 SIRRWFHGGRC 177

RESULT 13
Q9ZNB4_PORGI PRELIMINARY; PRT; 211 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DB 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Insertion sequence IS1126-like gene.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=381;
RX MEDLINE=95143166; PubMed=988746; DOI=10.1074/jbc.274.8.5012;
RA Shibata Y., Hayakawa M., Takiguchi H., Shiroza T., Abiko Y.;
RT "Determination and characterization of the hemagglutinin-associated
RT short motifs found in Porphyromonas gingivalis multiple gene
RT products.";
RL J. Biol. Chem. 274:5012-5020(1999).
DR EMBL; AB019363; BAA34342.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1-
DR PIR; P01609; Transposase_11; 1-
SQ SEQUENCE 211 AA; 24444 MW; 77040F65D76D1175 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLHGGFC 12
::|:|:|
Db 167 SIRRWFHGGRC 177

RESULT 14
Q9ZNB4_PORGI PRELIMINARY; PRT; 361 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DB 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=99150246; PubMed=10024556;
RA Hanley S.A., Aduse-Opoku J., Curtis M.A.;
RT "A 55-kilodalton immunodominant antigen of Porphyromonas gingivalis
RT W50 has arisen via horizontal gene transfer.";
RL Infect. Immun. 67:1157-1171(1999).
DR EMBL; AJ130872; CAA10225.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR PIR; P01609; Transposase_11; 1.
SQ SEQUENCE 361 AA; 41821 MW; F43664DC40E012C CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
```

```
Q7MUZ9_PORGI PRELIMINARY; PRT; 211 AA.
AC Q7MUZ9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DB 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ISPg1, transposase, internal deletion.
GN OrderedLocustNames=PQ1320;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017176; AAQ66391.1; -; Genomic_DNA.
DR TIGR; PG1320; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1-
DR Complete proteome.
SQ SEQUENCE 211 AA; 24300 MW; 86F8E3CAA3289A71 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 211;
Best Local Similarity 54.5%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLHGGFC 12
::|:|:|
Db 167 SIRRWFHGGRC 177

RESULT 14
Q9ZNB4_PORGI PRELIMINARY; PRT; 361 AA.
AC Q9ZNB4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DB 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=W50;
RX MEDLINE=99150246; PubMed=10024556;
RA Hanley S.A., Aduse-Opoku J., Curtis M.A.;
RT "A 55-kilodalton immunodominant antigen of Porphyromonas gingivalis
RT W50 has arisen via horizontal gene transfer.";
RL Infect. Immun. 67:1157-1171(1999).
DR EMBL; AJ130872; CAA10225.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR PIR; P01609; Transposase_11; 1.
SQ SEQUENCE 361 AA; 41821 MW; F43664DC40E012C CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
```

Mon May 15 11:35:36 2006

Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLGHC 12
Db 317 SRRWFHGGRC 327

RESULT 15

Q9ZAD0_PORGI PRELIMINARY; PRT; 361 AA.
ID Q9ZAD0_PORGI PRELIMINARY; PRT; 361 AA.
AC Q9ZAD0;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE ORF7.
GN Name=orf7;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=ATCC 33277;
RC MEDLINE=99198959; PubMed=10100860; DOI=10.1016/S0014-5793(99)00237-9;
RA Yoshida A., Nakano Y., Yamashita Y., Oho T., Shibata Y., Ohishi M.,
RT Koga T.;
RL FEBS Lett. 446:287-291(1999).
DR EMBL; AB015879; BAA35090.1; -; Genomic_DNA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR PIR; P01609; Transposase_11; 1.
SQ SEQUENCE 361 AA; 41588 MW; F47766D275D597F1 CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLGHC 12
Db 317 SRRWFHGGRC 327

RESULT 16

Q7M7B5_PORGI PRELIMINARY; PRT; 361 AA.
ID Q7M7B5_PORGI PRELIMINARY; PRT; 361 AA.
AC Q7M7B5;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
DE ISF91, transposase.
GN OrderedLocustNames=PG0184, PG0825, PG1031, PG1177, PG1197, PG1448,
PG1624, PG1906;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=W83;
RC MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AB017173; AAQ65657.1; -; Genomic DNA.
DR EMBL; AB017174; AAQ65741.1; -; Genomic DNA.
DR EMBL; AB017176; AAQ66328.1; -; Genomic_DNA.
DR TIGR; PG0460; -;
DR TIGR; PG0549; -;
DR TIGR; PG1244; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 361 AA; 41588 MW; 394DDC1E40F7B0E1 CRC64;

DR EMBL; AB017172; AAQ65419.1; -; Genomic_DNA.
DR EMBL; AB017174; AAQ65980.1; -; Genomic_DNA.
DR EMBL; AB017175; AAQ66150.1; -; Genomic_DNA.
DR EMBL; AB017176; AAQ66277.1; -; Genomic_DNA.
DR EMBL; AB017176; AAQ66289.1; -; Genomic_DNA.
DR EMBL; AB017177; AAQ66497.1; -; Genomic_DNA.
DR EMBL; AB017177; AAQ66552.1; -; Genomic_DNA.
DR EMBL; AB017178; AAQ66889.1; -; Genomic_DNA.
DR TIGR; PG0184; -;
DR TIGR; PG0825; -;
DR TIGR; PG1031; -;
DR TIGR; PG1177; -;
DR TIGR; PG1197; -;
DR TIGR; PG1448; -;
DR TIGR; PG1624; -;
DR TIGR; PG1906; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR Complete proteome.
SQ SEQUENCE 361 AA; 41754 MW; 8D6D7CB8D3104A2C CRC64;

Query Match 57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TLREWLGHC 12
Db 317 SRRWFHGGRC 327

RESULT 17

Q7M7E9_PORGI PRELIMINARY; PRT; 361 AA.
ID Q7M7E9_PORGI PRELIMINARY; PRT; 361 AA.
AC Q7M7E9;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-FEB-2005 (TREMELrel. 29, Last annotation update)
DE ISF91, transposase.
GN OrderedLocustNames=PG0460, PG0549, PG1244;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=W83;
RC MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AB017173; AAQ65657.1; -; Genomic DNA.
DR EMBL; AB017174; AAQ65741.1; -; Genomic DNA.
DR EMBL; AB017176; AAQ66328.1; -; Genomic_DNA.
DR TIGR; PG0460; -;
DR TIGR; PG0549; -;
DR TIGR; PG1244; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 361 AA; 41588 MW; 394DDC1E40F7B0E1 CRC64;

```

Query Match      57.7%; Score 45; DB 2; Length 361;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 TLRLWLGHC 12
Db      317 SRRWFGRC 327
      :|:|:|:|:|
      ||:|:|:|:|

RESULT 18
Q67UT0 ORYSA
ID Q67UT0 ORYSA PRELIMINARY; PRT; 396 AA.
AC Q67UT0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative serine/threonine protein kinase.
GN Name=P0046G12.38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota, Viridiplantae, Streptophyta, Tracheophyta;
OC Spermatophyta, Magnoliophyta, Liliopsida, Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
clone:P0046G12."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP005419; BADJ8089.1; -, Genomic_DNA.
DR Gramene; Q67UT0; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; PKinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS0107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS0108; PROTEIN KINASE_ST; 1.
DR ATP-binding; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 396 AA; 43831 MW; B2C9FB83CF34445E CRC64;

Query Match      57.7%; Score 45; DB 2; Length 396;
Best Local Similarity 77.8%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TLRLWLGHC 10
Db      197 TLQWLHGG 205
      ||:|:|:|:|
      ||:|:|:|:|

RESULT 19
Q9HRG8 HALSA
ID Q9HRG8 HALSA PRELIMINARY; PRT; 503 AA.
AC Q9HRG8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vng0705c.
GN OrderedLocusNames=VNG0705C;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.

```

```

OX NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 70022 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauser B., Kellier K., Cruz R., Danson M.J., Hough D.W., Dale H.,
RA Maddocks D.G., Jablonki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005016; AAG19190.1; -, Genomic_DNA.
DR PIR; B84228; B84228.
DR InterPro; IPR007357; DRRP.
DR Pfam; PF04244; DRRP; 1.
KW Complete proteome.
SQ SEQUENCE 503 AA; 56619 MW; 25DSACCFB886BC1A CRC64;

Query Match      57.7%; Score 45; DB 2; Length 503;
Best Local Similarity 77.8%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 LREWLHGGF 11
Db      379 LTFWFGHGF 387
      |:|:|:|:|
      |:|:|:|:|

RESULT 20
CBPY_YEAST
ID CBPY_YEAST STANDARD; PRT; 532 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
GN Name=PCY1; OrderedLocusNames=YMR297W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=87131100; PubMed=3028649; DOI=10.1016/0092-8674(87)90085-7;
RA Valls L.A., Hunter C.P., Rothman J.H., Stevens T.H.;
RT "Protein sorting in yeast: the localization determinant of yeast
vacuolar carboxypeptidase Y resides in the propeptide."
RL Cell 48:887-897(1987).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S288C / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93(1997).
RN [3]
RP PROTEIN SEQUENCE OF 112-532.
RA Svendsen I., Martin B.M., Viswanatha T., Johansen J.T.;
RT "Amino acid sequence of carboxypeptidase Y. II. Peptides from
enzymatic cleavages."
RL Carlsberg Res. Commun. 47:15-27(1982).
RN [4]
RP SEQUENCE REVISION, AND ACTIVE SITE SER-257.
RA Breddam K., Svendsen I.;
RT "Identification of methionyl and cysteinyl residues in the substrate
binding site of carboxypeptidase Y."

```


RL Carlsberg Res. Commun. 49:639-645(1984).
RN [5]
RP ACTIVE SITE HIS-508.
RX MEDLINE=90315013; PubMed=2639680;
RA Bech L.M., Breddam K.;
RT "Inactivation of carboxypeptidase Y by mutational removal of the
RT putative essential histidyl residue.";
RL Carlsberg Res. Commun. 54:165-171(1989).
RN [6]
RP MUTAGENESIS.
RX MEDLINE=94114535; PubMed=7904479;
RA Mortensen U.H., Remington S.J., Breddam K.;
RT "Site-directed mutagenesis on (serine) carboxypeptidase Y. A hydrogen
RT bond network stabilizes the transition state by interaction with the
RT C-terminal carboxylate group of the substrate.";
RL Biochemistry 33:508-517(1994).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=95244421; PubMed=7727362;
RA Endrizzi J.A., Breddam K., Remington S.J.;
RT "2.8-A structure of yeast serine carboxypeptidase.";
RL Biochemistry 33:1106-1120(1994).
CC -!- FUNCTION: Involved in degradation of small peptides. Digests
CC preferentially peptides containing an aliphatic or hydrophobic
CC residue in P1' position, as well as methionine, leucine or
CC phenylalanine in P1 position of ester substrate.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with broad
CC specificity.
CC -!- ENZYME REGULATION: Inhibited by ZPCK.
CC -!- SUBCELLULAR LOCATION: Lysosome-like vacuoles.
CC -!- PTM: Enters the endoplasmic reticulum as an inactive zymogen and
CC is modified by four N-linked core oligosaccharides, giving rise to
CC a precursor known as P1 (67 kDa). As P1 transits through the
CC Golgi, extension of its core oligosaccharides leads to the Golgi-
CC modified P2 precursor (69 kDa). P2 is sorted away from secretory
CC proteins at or beyond a late Golgi compartment and is subsequently
CC delivered to the vacuole via a prevacuolar endosome-like
CC compartment. Upon arrival in the vacuole, the N-terminal
CC prosegment of P2 is cleaved to yield the enzymatically active
CC mature vacuolar form of CPY (61 kDa).
CC -!- SIMILARITY: Belongs to the peptidase S10 family.
CC -!- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/COY/".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; M15482; AAA34902.1; -; Genomic_DNA.
CC EMBL; X80836; CAA56806.1; -; Genomic_DNA.
CC FJR; A26597; CPBY.
CC PDB; 1CPY; X-ray; @=112-532.
CC PDB; 1WPX; X-ray; A=112-532.
CC PDB; 1YSC; X-ray; @=112-532.
CC GerMOnline; 142974; -.
CC MEROPS; S10.001; -.
CC Ensembl; YMR297W; Saccharomyces cerevisiae.
CC SGD; S00004912; PRCL.
CC GO; GO:0005783; C:endoplasmic reticulum; IDA.
CC GO; GO:0000328; C:vacuolar lumen (sensu Fungi); TAS.
CC DR; GO:0004186; F:carboxypeptidase C activity; TAS.
CC DR; GO:0007039; P:vacuolar protein catabolism; TAS.
CC DR; InterPro; IPR008442; Carbp_N.
CC DR; InterPro; IPR001563; Peptidase_S10.
CC DR; InterPro; IPR000379; Ser_estra.
CC DR; PANTHER; PTHR11802; Peptidase_S10; 1.
CC DR; Pfam; PF05388; Carbp_N; 1.
CC DR; Pfam; PF00450; Peptidase_S10; 1.
CC DR; PRINTS; PR00724; CRBOX1PTASEC.
CC DR; ProDom; PD001189; Peptidase_S10; 1.
CC DR; PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW 3D-structure; Carboxypeptidase; Complete proteome;
KW Direct protein sequencing; Glycoprotein; Hydrolase; Protease; Signal;
KW Vacuole; Zymogen.
FT SIGNAL 1 20 Potential.
FT PROPEP 21 111
FT CHAIN 112 532 Carboxypeptidase Y.
FT MOTIF 24 27 Vacuolar targeting signal.
FT ACT_SITE 257 257 By similarity.
FT ACT_SITE 449 449
FT ACT_SITE 508 508 Substrate.
FT BINDING 452 452 Substrate.
FT BINDING 509 509 Substrate.
FT CARBOHYD 124 124 N-linked (GlcNAc. .).
FT CARBOHYD 198 198 N-linked (GlcNAc. .).
FT CARBOHYD 279 279 N-linked (GlcNAc. .).
FT CARBOHYD 479 479 N-linked (GlcNAc. .).
FT DISULFID 167 409
FT DISULFID 304 318
FT DISULFID 328 351
FT DISULFID 335 344
FT DISULFID 373 379
FT MUTAGEN 508 508
FT CONFLICT 260 261 H->A, R: Inactivates enzyme.
FT CONFLICT 389 389 GH -> HG (in Ref. 3).
FT CONFLICT 529 529 Y -> E (in Ref. 3).
FT STRAND 114 114 G -> D (in Ref. 3).
FT HELIX 116 118
FT STRAND 129 134
FT TURN 135 138
FT STRAND 139 146
FT TURN 152 154
FT STRAND 157 161
FT TURN 164 166
FT STRAND 167 167
FT TURN 169 169
FT HELIX 170 173
FT TURN 174 177
FT STRAND 180 183
FT TURN 184 186
FT STRAND 187 190
FT TURN 192 193
FT HELIX 195 198
FT STRAND 200 202
FT TURN 208 209
FT TURN 211 212
FT STRAND 214 215
FT HELIX 224 240
FT TURN 242 243
FT TURN 245 248
FT STRAND 251 256
FT TURN 257 258
FT HELIX 271 272
FT STRAND 282 286
FT HELIX 292 295
FT HELIX 296 298
FT HELIX 299 303
FT TURN 304 305
FT HELIX 315 338
FT TURN 341 355
FT TURN 356 356
FT HELIX 357 362
FT STRAND 365 365
FT TURN 366 367
FT STRAND 368 368
FT TURN 381 381
FT HELIX 382 391
FT HELIX 393 398
FT TURN 399 400
FT STRAND 409 409
FT HELIX 411 418
FT TURN 419 421

FT HELIX 422 424
 FT TURN 427 427
 FT HELIX 428 436
 FT TURN 437 438
 FT STRAND 441 446
 FT TURN 447 448
 FT TURN 451 452
 FT HELIX 454 463
 FT TURN 467 468
 FT STRAND 469 474
 FT TURN 478 481
 FT STRAND 483 485
 FT TURN 488 492
 FT STRAND 494 495

Query Match 57.7%; Score 45; DB 1; Length 532;
 Best Local Similarity 66.7%; Pred. No. 90;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LREWLHGPF 11
 : ||: |||

Db 522 VNEIHHGPF 530

RESULT 21

ID Q7V2M4 PROMP PRELIMINARY; PRT; 154 AA.

AC Q7V2M4;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN OrderedLocustNames=PM0579;

OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).

OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcales;

OC Prochlorococcus.

OX NCBI_TaxID=59919;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;

RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,

RA Ahlgren N.A., Atellano A., Coleman M., Hauser L., Hess W.R.,

RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,

RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,

RA Webb E.A., Zinser E.R., Chisholm S.W.;

RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic

RT niche differentiation."

RL Nature 424:1042-1047(2003).

DR EMBL; BX572091; CAB19038.1; -; Genomic_DNA.

DR InterPro; IPR007115; 6_PTP_synth.

DR PIRSF; PIRSF006113; PTP_syn; 1.

DR ProDom; PD004049; PTPS_hypoth; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 154 AA; 18598 MW; D5E9B5EE909C89BC CRC64;

Query Match 56.4%; Score 44; DB 2; Length 154;

Best Local Similarity 50.0%; Pred. No. 37;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTRLEWLHGPF 12
 : ||: |||

Db 19 CSHRQWRHDGHC 30

RESULT 22

ID Q9ZPU0 ARATH PRELIMINARY; PRT; 171 AA.

AC Q9ZPU0;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein At2g13980.

GN Name=At2g13980;

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.B., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC006528; AAD19777.1; -; Genomic_DNA.

DR PIR; G84512; G84512.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0004518; F:nuclease activity; IEA.

DR InterPro; IPR012337; RNaseH_fold.

KW Hydrolase; Hypothetical protein; Nuclease.

SQ SEQUENCE 171 AA; 19347 MW; D615373DE0453A5A CRC64;

Query Match 56.4%; Score 44; DB 2; Length 171;

Best Local Similarity 58.3%; Pred. No. 42;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTRLEWLHGPF 12
 : ||: |||

Db 156 CTRLEWLHGPF 167

RESULT 23

ID Q6CBE1 YARLI PRELIMINARY; PRT; 539 AA.

AC Q6CBE1;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Similar to CA0795|IPF15641 Candida albicans IPF15641 unknown

DE function.

GN OrderedLocustNames=YALI0C19682g;

OS Yarrowia lipolytica (Candida lipolytica).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Dipodascaceae; Yarrowia.

OX NCBI_TaxID=4952;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX PubMed=15229592; DOI=10.1038/nature02579;

RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,

RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,

RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,

RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,

RA Deapans L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,

RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,

RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller R.,

RA Nicoud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,

RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,

RA Swennen D., Tekala P., Wesolowski-Louvel M., Westhof E., Wirth B.,

RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,

RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,

RA Wincker P., Souciet J.-L.;

RT "Genome evolution in yeasts."

RL Nature 430:35-44(2004).

DR EMBL; CR382129; CAG82341.1; -; Genomic_DNA.

DR InterPro; IPR003347; PF_JmjC.

DR Pfam; PF02373; JmjC; 1.

DR SMART; SM00558; JmjC; 1.

KW Complete proteome.

SQ SEQUENCE 539 AA; 61646 MW; AF9764FD74AEB954 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 539;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
|:|:|:|:|:
Db 376 CSLAEWPLGGF 386

RESULT 24
Q61MS4 CAEBR
ID Q61MS4 CAEBR PRELIMINARY; PRT; 731 AA.
AC Q61MS4;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG08359.
GN Name=CBG08359;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
family.
DR EMBL; CAAC0100035; CAE63810.1; -; Genomic_DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR PRINTS; PR00154; AMPBINDING
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Hypothetical protein.
SQ SEQUENCE 731 AA; 81445 MW; C071504A05B95631 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 731;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LREWLHGGF 11
|:|:|:|:|:
Db 517 LREWLEGGY 525

RESULT 25
Q628B7 CAEBR
ID Q628B7 CAEBR PRELIMINARY; PRT; 1540 AA.
AC Q628B7;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG00449.
GN Name=CBG00449;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG NIH - Mammalian Gene Collection (MGC) project;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; CAAC0100004; CAE57481.1; -; Genomic_DNA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.

DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 3.
KW EGF-like domain; Hypothetical protein.
SQ SEQUENCE 1540 AA; 168546 MW; 48CECC95069F6E60 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 1540;
Best Local Similarity 63.6%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTLREWLHGGF 11
|:|:|:|:|:
Db 35 CTIRECLNGGY 45

RESULT 26
Q61EJ2 CAEBR
ID Q61EJ2 CAEBR PRELIMINARY; PRT; 4856 AA.
AC Q61EJ2;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG12050.
GN Name=CBG12050;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; CAAC0100059; CAE66705.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002172; LDL receptor A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF00057; Ldl_recept_a; 28.
DR Pfam; PF00058; Ldl_recept_b; 18.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 25.
DR SMART; SM00179; EGF_CA; 8.
DR SMART; SM00192; LDLA; 32.
DR SMART; SM00135; LY; 27.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS50026; EGF_3; 9.
DR PROSITE; PS01209; LDLRA_1; 18.
DR PROSITE; PS50068; LDLRA_2; 24.
KW Hypothetical protein.
SQ SEQUENCE 4856 AA; 541409 MW; D10DA26DC9A5CC8A CRC64;

Query Match 56.4%; Score 44; DB 2; Length 4856;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTLREWLHGGFC 12
|:|:|:|:|:
Db 4464 CKDRECLNGGFC 4475

RESULT 27

```
Q7U636 SYNXP
ID Q7U636 SYNXP PRELIMINARY; PRT; 158 AA.
AC Q7U636
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SYNW1504;
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regalia W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569693; CAB08019.1; -; Genomic_DNA.
DR InterPro; IPR007115; 6 PTP_synth.
DR PIRSF; PIRSF006113; PTP_syn; 1.
DR ProDom; PD004049; PTPS_Hypoth; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 158 AA; 18232 MW; DB0A4EDAD8831AAB CRC64;

Query Match 55.1%; Score 43; DB 2; Length 158;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CTLREWLHGFGC 12
| | | | |
Db 19 CCHQWRHAGC 30

RESULT 28
Q9B598 CHILA
ID Q9B598 CHILA PRELIMINARY; PRT; 182 AA.
AC Q9B598
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=Cytb;
OS Chinchilla lanigera (Chinchilla).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
OC Hystricognathi; Chinchillidae; Chinchilla.
OX NCBI_TaxID=34839;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Skin;
RA Spotorno A.E., Valladares J.P., Zuleta C., Marin J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COPACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF244380; AAK28364.1; -; Genomic_DNA.
DR SMR; Q9B598; 13-182.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
```

```
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 182
SQ SEQUENCE 182 AA; 20594 MW; CAA7A70BFF51374E CRC64;

Query Match 55.1%; Score 43; DB 2; Length 182;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGFG 11
| | | | |
Db 163 TLEWIWGGF 172

RESULT 29
Q5V5F2 HALMA
ID Q5V5F2 HALMA PRELIMINARY; PRT; 328 AA.
AC Q5V5F2
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative transporter.
GN OrderedLocNames=trnAC0187;
OS Haloarcula marismortui (Halo bacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=2238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43049;
RX PubMed=15520287; DOI=10.1101/gr.2700304;
RA Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G.,
RA Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P.,
RA Date S.V., Marcotte E., Hood L., Ng W.V.;
RT "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
RT the Dead Sea.";
RL Genome Res. 14:2221-2234(2004).
DR EMBL; AY596297; AAV45250.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 328 AA; 35627 MW; 4DC488FB91107046 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 328;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TLREWLHGFG 11
| | | | |
Db 293 TLEWIWGGW 302

RESULT 30
Q9B640 GRODE
ID Q9B640 GRODE PRELIMINARY; PRT; 380 AA.
AC Q9B640
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=Cytb;
OS Eothenomys andersoni.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetidae; Arvicolinae; Eothenomys.
OX NCBI_TaxID=82466;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Iwasa M.A., Suzuki H.;
RT "Evolutionary networks of maternal and paternal gene lineages in voles
RT (Eothenomys) endemic to Japan.";
```

RL J. Mammal. 83:852-865(2002).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex III or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AB037303; BAB40558.1; -; Genomic_DNA.
 DR SMR; Q9B640; 2-377.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; F: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.
 DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; Cytochrom_B_C; 1.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS10033; CYTB_CTER; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 380 380
 SQ SEQUENCE 380 AA; 42882 MW; 188E774B649B44BD CRC64;

 Query Match 55.1%; Score 43; DB 2; Length 380;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 2 TLREWLGHGF 11
 DB 159 TLEEWGWF 168

 RESULT 31
 ID Q8WF54_9BIVA PRELIMINARY; PRT; 440 AA.
 AC Q8WF54;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b.
 GN Name-cytb;
 OS Venerupis (Ruditapes) philippinarum.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
 OC Veneroidea; Veneridae; Venerupis; Ruditapes.
 OX NCBI_TaxID=129788;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Gonad;
 RA Okazaki M., Ueshima R.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex III or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 CC EMBL; AB065374; BAB83775.1; -; Genomic_DNA.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; F: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.

DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR Pfam; PF00032; Cytochrom_B_C; 1.
 DR Pfam; PF00033; Cytochrom_B_N; 1.
 DR PROSITE; PS10033; CYTB_CTER; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 SQ SEQUENCE 440 AA; 49604 MW; 93C3A7678C8CE2FD CRC64;

 Query Match 55.1%; Score 43; DB 2; Length 440;
 Best Local Similarity 63.6%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

 QY 2 TLREWLGHGF 12
 DB 170 TLEWVWGGHC 180

 RESULT 32
 ID Q4ICS6_GIBZE PRELIMINARY; PRT; 558 AA.
 AC Q4ICS6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=FG04982.1;
 OS Gibberella zeae PH-1.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
 OX NCBI_TaxID=229533;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PH-1;
 RA Arrachchi H.M., Barna N., Nusbbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Boukhgaltier B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil S.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severi P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Willson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RL "Fusarium graminearum genome sequence.";
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL; AACM01000200; BAA74849.1; -; Genomic_DNA.
 DR Hypothetical protein.
 SQ SEQUENCE 558 AA; 61297 MW; FACB1CEFF89AD96 CRC64;

 Query Match 55.1%; Score 43; DB 2; Length 558;
 Best Local Similarity 66.7%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

 QY 2 TLREWLGHGF 10

```
Db 107 SMRWLHGG 115
:|:|:|:|
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

RESULT 33
Q8CY51_BIFLO
ID Q8CY51_BIFLO PRELIMINARY; PRT; 583 AA.
AC Q8CY51_BIFLO PRELIMINARY; PRT; 583 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=BL1010;
OS Bifidobacterium longum
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schell M.A., Karmirantzou M., Snel B., Villanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).
DR EMBL; AE014295; AAN4818.1; -; Genomic DNA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR002543; FtsK SpoIIIE.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS05901; FTSK; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 583 AA; 62966 MW; 6487ED8EDC338CA0 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 583;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 BMLHGGFC 12
:|:|:|:|
Db 459 QWLHGGRC 466

RESULT 34
Q59GS6_HUMAN
ID Q59GS6_HUMAN PRELIMINARY; PRT; 605 AA.
AC Q59GS6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Diacylglycerol kinase, gamma 90kDa variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Torok Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB209033; BAD22270.1; -; mRNA.
KW Kinase.
FT NON TER
SQ SEQUENCE 605 AA; 68332 MW; 18E1846CA4366F2E CRC64;

Query Match 55.1%; Score 43; DB 2; Length 605;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TLREWHLGG 10
:|:|:|:|
Db 244 SLQEWVHGG 252

RESULT 35
Q5FWGI_HUMAN
ID Q5FWGI_HUMAN PRELIMINARY; PRT; 766 AA.
AC Q5FWGI;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DGKG protein.
GN Name=DGKG;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymph;
RA Director MGC Project;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC089411; AAH89411.1; -; mRNA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004143; F:diacylglycerol kinase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0007205; P:protein kinase C activation; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000756; DAGK.
DR InterPro; IPR001206; DAGK.
DR InterPro; IPR002048; EF_hand_Ca_bd.
DR Pfam; PF00130; C1_1; 2.
DR Pfam; PF00609; DAGK_acc; 1.
DR Pfam; PF00781; DAGK_cat; 1.
DR Pfam; PF00036; ehand; 2.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR ProDom; PD002939; DAGK; 1.
DR ProDom; PD005043; DAGK; 1.
DR ProDom; PD00012; EF-hand; 1.
DR SMART; SM00109; C1; 2.
DR SMART; SM00045; DAGK; 1.
DR SMART; SM00046; DAGK; 1.
DR SMART; SM00054; EPH; 2.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
```

```
DR PROSITE, PS50081; DAG PE BIND_DOM_2; 2.
DR PROSITE, PS00018; EF_HAND; UNKNOWN_2.
KW Calcium; Repeat.
SQ SEQUENCE 766 AA; 86303 MW; 1D3F8C78BD59219F CRC64;

Query Match      55.1%; Score 43; DB 2; Length 766;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 TLREWLGHG 10
   :|||:|
Db 241 SLQEWVHG 249

RESULT 36
DGKG HUMAN STANDARD; PRT; 791 AA.
AC P49619;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Diacylglycerol kinase, gamma (SC 2.7.1.107) (Diglyceride kinase) (DGK-
DE gamma) (DAG kinase gamma).
GN Name=DGKG; Synonyms=DAGK3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=Liver;
RC MEDLINE=94308084; PubMed=8034597;
RA Kai M., Sakane F., Inai S.-I., Wada I., Kanoh H.;
RT "Molecular cloning of a diacylglycerol kinase isozyme predominantly
RT expressed in human retina with a truncated and inactive enzyme
RT expression in most other human cells.";
RL J. Biol. Chem. 269:18492-18498(1994).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99168758; PubMed=10071200; DOI=10.1007/s004390050917;
RA Stoeckl H., Klein J., Gehrig A., Koehler M.R., Jurkies B., Kellner U.,
RA Leo-Kottler B., Schmid M., Weber B.H.F.;
RT "Mapping and genomic characterization of the gene encoding
RT diacylglycerol kinase gamma (DAGK3): assessment of its role in
RT dominant optic atrophy (OPA1).";
RL Hum. Genet. 104:99-105(1999).
CC -!- FUNCTION: Reverses the normal flow of glycerolipid biosynthesis by
CC phosphorylating diacylglycerol back to phosphatidic acid.
CC -!- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-diacyl-
CC sn-glycerol 3-phosphate.
CC -!- ENZYME REGULATION: Requires phosphatidylserine for maximal
CC activity.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Can be loosely bound to the
CC membranes.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P49619-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P49619-2; Sequence=VSP_001267;
CC Note=May be inactive;
CC -!- TISSUE SPECIFICITY: Predominantly expressed in retina and in a
CC much lesser extent in the brain. Other tissues contain extremely
CC low levels of DGK-gamma.
CC -!- SIMILARITY: Belongs to the eukaryotic diacylglycerol kinase
CC family.
CC -!- SIMILARITY: Contains 2 EF-hand domains.
CC -!- SIMILARITY: Contains 2 phorbol-ester/DAG-type zinc fingers.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
```

use as long as its content is in no way modified and this statement is not removed.

CC EMBL; D26135; BAA05132.1; -; mRNA.

DR EMBL; AF020945; AAC04686.1; -; Genomic DNA.

DR EMBL; AF020922; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020923; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020924; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020925; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020926; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020927; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020928; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020929; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020930; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020931; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020932; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020933; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020934; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020935; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020936; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020937; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020938; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020939; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020940; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020941; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020942; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020943; AAC04686.1; JOINED; Genomic DNA.

DR EMBL; AF020944; AAC04686.1; JOINED; Genomic DNA.

DR PIR; A53691; A53691.

DR Ensembl; ENSG0000058866; Homo sapiens.

DR HGNC; HGNC:2853; DGKG.

DR MIM; 601854; .

DR GO; GO:0004143; F:diacylglycerol kinase activity; TAS.

DR GO; GO:0007165; P:signal transduction; TAS.

DR InterPro; IPR002219; DAG_PE_bd.

DR InterPro; IPR000756; DAGKa.

DR InterPro; IPR001206; DAGKc.

DR InterPro; IPR011992; EF-Hand type.

DR InterPro; IPR002048; EF_hand_Ca_bd.

DR Pfam; PF00130; C1_1; 2.

DR Pfam; PF06509; DAGK_acc; 1.

DR Pfam; PF00781; DAGK_cat; 1.

DR Pfam; PF00036; ehand; 2.

DR PRINTS; PR000008; DAGPEDOMAIN.

DR ProDom; PD002939; DAGKa; 1.

DR ProDom; PD005043; DAGKc; 1.

DR ProDom; PD000012; EF-hand; 1.

DR PROSITE; PS00018; EF_HAND_1; 2.

DR PROSITE; PS02222; EF_HAND_2; 2.

DR PROSITE; PS00479; ZF_DAG_PE_1; 2.

DR PROSITE; PS00081; ZF_DAG_PE_2; 2.

KW Alternative splicing; Calcium; Kinase; Metal-binding;

KW Multigene family; phorbol-ester binding; Polymorphism; Repeat;

KW Transferase; Zinc; Zinc-finger.

FT DOMAIN 175 210 EF-hand 1.

FT DOMAIN 220 255 EF-hand 2.

FT CA_BIND 188 199 1 (Potential).

FT CA_BIND 233 244 2 (Potential).

FT ZN_FING 271 321 Phorbol-ester/DAG-type 1.

FT ZN_FING 336 385 Phorbol-ester/DAG-type 2.

FT REGION 432 558 Catalytic-A (Potential).

FT REGION 578 752 Catalytic-B (Potential).

FT COMPBIAS 151 156 Poly-Ser.

FT VARSPLIC 451 475 Missing (in isoform Short).

FT VARIANT 142 142 T -> S (in dbSNP:1004588).

FT VARIANT 370 370 R -> W (in dbSNP:3213770).

FT VARSPLIC 791 791 C7DD07F5B285FF62 CRC64;

SQ SEQUENCE 791 AA; 88997 MW; C7DD07F5B285FF62 CRC64;

Query Match 55.1%; Score 43; DB 1; Length 791;

Best Local Similarity 66.7%; Pred. No. 2.9e+02;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TIREWLHGG 10
Db 241 SLQEWVHGG 249

RESULT 17

Q53JX1 ORYSA
ID Q53JX1 ORYSA PRELIMINARY; PRT; 884 AA.
AC Q53JX1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Retrotransposon protein, putative, unclassified.
GN ORName=LOC_Os11g22850;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,
RA Bera J., Kim M., Jin S., Padrosh D., Vuong H., Overton II L., Aken S.,
RA Rardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,
RA de Vazilhes A., White O., Salzberg S., Fraser C.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC146938; AAX96001.1; -; Genomic DNA.
SQ SEQUENCE 884 AA; 98530 MW; A3E9533439EDE3B1 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 884;
Best Local Similarity 55.6%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 REWLHGGFC 12
Db 191 RDWHANFC 189

RESULT 38

Q9D245 MOUSE
ID Q9D245 MOUSE PRELIMINARY; PRT; 104 AA.
AC Q9D245;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Mus musculus adult male urinary bladder cDNA, RIKEN
DE enriched library, clone:9530039B02 product:hypothetical protein, full
DE insert sequence (Mus musculus 10 days neonate cerebellum cDNA, RIKEN
DE full-length enriched library, clone:B930046021 product:hypothetical
DE protein, full insert sequence).
GN Name=Cstadi;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Urinary bladder;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Urinary bladder;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaudo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedziera R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Negashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Urinary bladder;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Urinary bladder;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Mus musculus adult male aorta and vein cDNA, RIKEN full-length
DE enriched library, clone:A53008E08 product:hypothetical protein, full
DE insert sequence (Mus musculus 4 days neonate male adipose cDNA, RIKEN
DE full-length enriched library, clone:B43004P05 product:hypothetical
DE protein, full insert sequence).
GN Name=A430107D22Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nitaiko I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami H., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Hammond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasaki A., Kedzierski R.M., King B.L.,
RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nageshima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahrstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,

RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Adipose, and Aorta and vein;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tegawa A., Takahashi F., Takaku-Akai H., Muramatsu M., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK041179; BAC30852.1; -; mRNA.
DR EMBL; AK046549; BAC32781.1; -; Mus musculus.
DR Ensemble; ENSMUSG00000053467; Mus musculus.
DR MGI; MGI:2444128; A430107D22Rik.
KW Hypothetical protein.
SQ SEQUENCE 121 AA; 13068 MW; DA6405F40B706FEE CRC64;

Query Match 53.8%; Score 42; DB 2; Length 121;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 TLREWLHGG 10
||:|:|:|:
DB 47 TLQRVHGG 55

RESULT 42

P92837_9HYME
ID P92837_9HYME PRELIMINARY; PRT; 128 AA.
AC P92837;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Ephedrus persicae.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonidea;
OC Braconidae; Aphidinae; Ephedrus.

```

OX NCBI_TaxID=55917;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97330629; PubMed=9187088; DOI=10.1006/mpev.1996.0400;
RA Beishaw R.D., Quicke D.L.J.;
RT "A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae)."
RL Mol. Phylogenet. Evol. 7:281-293(1997).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; Z83630; CAB05972.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS10033; Cytochrom_B_N; 1.
DR PROSITE; PS1002; CYTB_CTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 15261 MW; 62009B9B3D42E3C5 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 128;
Best Local Similarity 80.0%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLGGF 11
DB 12 TLVEWLGGF 21

RESULT 43
Q7YAZ1_9SAUR PRELIMINARY; PRT; 134 AA.
AC Q7YAZ1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name-cytc;
OS Colobosaura modesta.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Teliioidea;
OC Gymnophthalmidae; Colobosaura.
OX NCBI_TaxID=174786;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22980429; PubMed=14615195; DOI=10.1016/S1055-7903(03)00142-8;
RA Whiting A.S., Bauer A.M., Sites J.W. Jr.;
RT "Phylogenetic relationships and limb loss in sub-Saharan African
RT scincine lizards (Squamata: Scincidae).";
RL Mol. Phylogenet. Evol. 29:582-598(2003).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,

```

```

CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AY217799; AAP45101.1; -; Genomic_DNA.
DR SMR; Q7YAZ1; 1-134.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS10033; CYTB_CTER; 1.
DR PROSITE; PS1002; CYTB_NTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15106 MW; 97C12F6552055E8 CRC64;

Query Match 53.8%; Score 42; DB 2; Length 134;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLGGF 11
DB 23 TLVEWLGGF 32

RESULT 44
Q9MMC3_GUIGU PRELIMINARY; PRT; 144 AA.
AC Q9MMC3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Guira guira (Guira cuckoo).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Cuculiformes; Crotophagidae; Guira.
OX NCBI_TaxID=30392;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20179528; PubMed=10712848; DOI=10.1006/mpev.1999.0715;
RA Johnson K.P., Goodman S.M., Lanyon S.M.;
RT "A phylogenetic study of the malagasy couas with insights into cuckoo
RT relationships."
RL Mol. Phylogenet. Evol. 14:436-444(2000).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; AF204998; AAF68885.1; -; Genomic_DNA.
DR SMR; Q9MMC3; 1-144.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS1002; CYTB_NTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.

```

```

FT  NON TER      1      1
FT  NON TER     144     144
SQ  SEQUENCE    144 AA, 16103 MW, 2FCE9A0B86DD39C9 CRC64;

Query Match      53.8%; Score 42; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. NO. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 TLREWLGGF 11
    |||||
Db  127 TLVEWLGGF 136

RESULT 45
ID  Q9MMC4_9AVES PRELIMINARY; PRT; 144 AA.
AC  Q9MMC4;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Crotophaga major.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Cuculiformes; Crotophagidae;
OC  Crotophaga.
OX  NCBI_TaxID=48617;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RX  MEDLINE=20179528; PubMed=10712848; DOI=10.1006/mpev.1999.0715;
RA  Johnson K.P., Goodman S.M., Lanyon S.M.;
RT  "A phylogenetic study of the malgasy couas with insights into cuckoo
RL  Mol. Phylogenet. Evol. 14:436-444 (2000).
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-cl complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AF204997; AAF68884.1; -; Genomic_DNA.
DR  SMR; Q9MMC4; 1-144.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016020; C:membrane; IEA.
DR  GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR  GO; GO:0005739; C:mitochondrion; IEA.
DR  GO; GO:0016431; F:oxidoreductase activity; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR005797; Cytb_b6_N.
DR  Pfam; PF00033; Cytochrom B_N; 1.
DR  PROSITE; PS51002; Cytb_NTER; 1.
KW  Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW  Respiratory chain; Transmembrane; Transport.
FT  NON TER      1      1
FT  NON TER     144     144
SQ  SEQUENCE    144 AA, 16111 MW, 5423190720CCFFA4 CRC64;

Query Match      53.8%; Score 42; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. NO. 74;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 TLREWLGGF 11
    |||||
Db  127 TLVEWLGGF 136

RESULT 46
Q4VTU6_9SAUR PRELIMINARY; PRT; 193 AA.
ID  Q4VTU6_9SAUR PRELIMINARY;
AC  Q4VTU6;

```

```

DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT  13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Gymnodactylus darwinii.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
OC  Gymnodactylus.
OX  NCBI_TaxID=298113;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=MRTA2;
RA  Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
RA  Yasuda Y.Y., Sites J.W. Jr.;
RT  "Phylogeography and species limits in the Gymnodactylus darwinii
RT  complex (Gekkonidae, Squamata): genetic structure coincides with river
RT  systems in the Brazilian Atlantic Forest.";
RL  Biol. J. Linn. Soc. Lond. 85:13-26(2005).
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-cl complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AY630394; AAV54568.1; -; Genomic_DNA.
DR  InterPro; IPR005797; Cytb_b6_N.
DR  Pfam; PF00033; Cytochrom B_N; 1.
DR  PROSITE; PS51002; Cytb_NTER; 1.
KW  Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW  Respiratory chain; Transmembrane; Transport.
FT  NON TER      1      1
FT  NON TER     193     193
SQ  SEQUENCE    193 AA, 21639 MW, CEEA98C5C321D19C CRC64;

Query Match      53.8%; Score 42; DB 2; Length 193;
Best Local Similarity 80.0%; Pred. NO. 1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  2 TLREWLGGF 11
    |||||
Db  141 TLVEWLGGF 150

RESULT 47
Q7YB81_9SYLV PRELIMINARY; PRT; 196 AA.
ID  Q7YB81_9SYLV PRELIMINARY;
AC  Q7YB81;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Cytochrome b (Fragment).
OS  Alcippe morrissonia.
OG  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; Alcippe.
OX  NCBI_TaxID=204791;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Pasquet E., Cibois A., Kalyakin M.V., Bourdon E.;
RL  Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC  complex (complex III or cytochrome b-cl complex), which is a
CC  respiratory chain that generates an electrochemical potential
CC  coupled to ATP synthesis (By similarity).
CC  -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC  -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CC  cytochrome c1 and the Rieske protein (By similarity).
CC  -!- SIMILARITY: Belongs to the cytochrome b family.
DR  EMBL; AY128557; AAN01293.1; -; Genomic_DNA.
DR  SMR; Q7YB81; 1-196.

```

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytochrome b6 C.
 DR InterPro; IPR005797; Cytochrome b6 N.
 DR Pfam; PF00032; Cytochrome b6 C; 1.
 DR PROSITE; PS11003; CYTB_CTER; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 196
 SQ SEQUENCE 196 AA; 21764 MW; A8476B4C2F3447E8 CRC64;
 Query Match 53.8%; Score 42; DB 2; Length 196;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGKF 11
 || || || || ||
 Db 22 TLVEWLWGGF 31

RESULT 48
 Q4VTU5_9SAUR PRELIMINARY; PRT; 203 AA.
 AC Q4VTU5;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Cytochrome b (Fragment).
 OS Gymnodactylus darwini.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
 OC Gymnodactylus.
 NCBI_TaxID=298113;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=MRTA3;
 RA Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
 RA Yassuda Y.Y., Sites J.W. Jr.;
 RT "Phylogeography and species limits in the Gymnodactylus darwini
 complex (Gekkonidae, Squamata): genetic structure coincides with river
 systems in the Brazilian Atlantic Forest.";
 RL Biol. J. Linn. Soc. Lond. 85:13-26(2005).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AY630396; AAV54569.1; -; Genomic_DNA.
 DR EMBL; AY630393; AAV54567.1; -; Genomic_DNA.
 DR InterPro; IPR005798; Cytochrome b6 C.
 DR InterPro; IPR005797; Cytochrome b6 N.
 DR Pfam; PF00033; Cytochrome b6 C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 203
 SQ SEQUENCE 203 AA; 22852 MW; F4FC8230EC586FCA CRC64;
 Query Match 53.8%; Score 42; DB 2; Length 203;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGKF 11
 || || || || ||
 Db 136 TLVEWLWGGF 145

RESULT 49
 Q4VTU4_9SAUR PRELIMINARY; PRT; 213 AA.
 AC Q4VTU4;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Cytochrome b (Fragment).
 OS Gymnodactylus darwini.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae;
 OC Gymnodactylus.
 NCBI_TaxID=298113;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=MRTA5 and MRTAL;
 RA Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M.,
 RA Yassuda Y.Y., Sites J.W. Jr.;
 RT "Phylogeography and species limits in the Gymnodactylus darwini
 complex (Gekkonidae, Squamata): genetic structure coincides with river
 systems in the Brazilian Atlantic Forest.";
 RL Biol. J. Linn. Soc. Lond. 85:13-26(2005).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 complex (complex III or cytochrome b-c1 complex), which is a
 respiratory chain that generates an electrochemical potential
 coupled to ATP synthesis (By similarity).
 CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AY630396; AAV54570.1; -; Genomic_DNA.
 DR EMBL; AY630393; AAV54567.1; -; Genomic_DNA.
 DR InterPro; IPR005798; Cytochrome b6 C.
 DR InterPro; IPR005797; Cytochrome b6 N.
 DR Pfam; PF00033; Cytochrome b6 C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 213
 SQ SEQUENCE 213 AA; 23895 MW; 2B6EA26AF385F7DD CRC64;
 Query Match 53.8%; Score 42; DB 2; Length 213;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TLREWLHGKF 11
 || || || || ||
 Db 136 TLVEWLWGGF 145

RESULT 50
 Q34157_CORMC PRELIMINARY; PRT; 215 AA.
 AC Q34157;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE Cytochrome b (Fragment).
 OS Corvus macrorhynchos (Jungle crow).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;
 OC Corvinidae; Corvus.
 NCBI_TaxID=36249;
 RN [1]

Mon May 15 11:35:36 2006

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Muscle;
RA Chikuni K., Minaka N., Ikenaga H.;
RT "Molecular phylogeny of some Passeriformes, based on cytochrome b
RL Yamashina Choru Kenkyujo Kenkyu Hokoku 28:1-8(1995).
CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -1- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome b family.
DR EMBL; D38313; BAA07429.1; -; Genomic_DNA.
DR SMR; Q34157; 1-215.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00033; Cytochrom_B_N_1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 215
SQ SEQUENCE 215 AA; 24057 MW; 65D3C67C2F496D1D CRC64;

Query Match 53.8%; Score 42; DB 2; Length 215;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TLREWLHGGF 11
Db 107 TLVEWLWGGF 116

```

Search completed: May 12, 2006, 10:49:09
 Job time : 99.9231 secs

OM protein - protein search, using sw model
 Run on: May 12, 2006, 10:19:39 ; Search time 101.709 Seconds
 (without alignments)
 60.479 Million cell updates/sec

 Title: US-10-632-388-53
 Perfect score: 81
 Sequence: 1 CSLQFLSHGGVVC 14

 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
 Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 1000 summaries

Database : A_Geneseq_21.*

1: Geneseqp1980s.*
 2: Geneseqp1990s.*
 3: Geneseqp2000s.*
 4: Geneseqp2001s.*
 5: Geneseqp2002s.*
 6: Geneseqp2003as.*
 7: Geneseqp2003bs.*
 8: Geneseqp2004s.*
 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	14	AAW09535	Aaw09535 Thrombopo
2	81	100.0	14	AAW36686	Aaw36686 Thrombopo
3	81	100.0	14	AAW16997	Aaw16997 TPO-mimet
4	81	100.0	14	AAU25905	Aau25905 Human thr
5	81	100.0	14	ABB72883	Abb72883 TPO mimet
6	81	100.0	14	ADJ73034	Adj73034 TPO mimet
7	81	100.0	14	ADJ52669	Adj52669 CHI delet
8	81	100.0	14	ADJ51630	Adj51630 CHI delet
9	49	60.5	453	ADT60670	Adt60670 Plant pol
10	48	59.3	344	AAW96482	Aaw96482 Human top
11	48	59.3	660	AAW81503	Aaw81503 Dead Box
12	48	59.3	660	ABO07237	Ab007237 Human p53
13	48	59.3	660	ADL13058	Adl13058 Human ste
14	48	59.3	660	ADM33393	Adm33393 Human PRO
15	48	59.3	660	ADW08765	Adw08765 Human pro
16	48	59.3	660	AEA24054	Aea24054 Human PRO
17	48	59.3	662	AAW81501	Aaw81501 Dead Box
18	48	59.3	662	AAW81502	Aaw81502 Dead Box
19	48	59.3	662	AAW96483	Aaw96483 Human BEX
20	48	59.3	662	ABB57045	Abb57045 Mouse isc
21	48	59.3	662	ABO07236	Ab007236 Human p53
22	48	59.3	662	ABO52958	Ab052958 Human spl
23	48	59.3	662	ADN04580	Adn04580 Antipsori
24	48	59.3	662	ADN03714	Adn03714 Antipsori

25	48	59.3	662	9	ADX07534	Adx07534 Cyclin-de
26	45	55.6	192	3	AAQ07259	Aaq07259 Arabidops
27	45	55.6	198	8	ADJ32021	Adj32021 Plant ino
28	45	55.6	251	3	AAQ05910	Aaq05910 Arabidops
29	45	55.6	313	3	AAV59428	Aav59428 Soybean i
30	45	55.6	370	3	AAQ05909	Aaq05909 Arabidops
31	45	55.6	536	6	ABU27233	Abu27233 Protein e
32	44	54.3	179	4	AAU273148	Aau273148 Novel hum
33	44	54.3	179	4	AAU87567	Aau87567 Novel hum
34	44	54.3	179	8	ADI54882	Adi54882 Novel hum
35	44	54.3	184	4	ABG26613	Abg26613 Novel hum
36	44	54.3	187	4	ABO3078	Ab03078 Human exp
37	44	54.3	187	4	AAU87270	Aau87270 Novel cen
38	44	54.3	187	8	ADI54585	Adi54585 Novel hum
39	44	54.3	196	3	AAQ44384	Aaq44384 Arabidops
40	44	54.3	222	3	AAQ44383	Aaq44383 Arabidops
41	44	54.3	238	8	ADJ32010	Adj32010 Plant ino
42	44	54.3	238	8	ADJ32009	Adj32009 Plant ino
43	44	54.3	247	8	ADJ32014	Adj32014 Plant ino
44	44	54.3	261	8	ADJ32016	Adj32016 Plant ino
45	44	54.3	279	8	ADJ32018	Adj32018 Arabidops
46	44	54.3	319	3	AAQ44382	Aaq44382 Arabidops
47	44	54.3	321	3	AAV59421	Aav59421 Corn inos
48	44	54.3	341	3	AAV59423	Aav59423 Rice inos
49	44	54.3	342	7	ADC13500	Adc13500 Protein w
50	44	54.3	352	8	ADT59474	Adt59474 Plant pol
51	44	54.3	399	3	AAV59425	Aav59425 Soybean i
52	44	54.3	431	6	ABM67394	Abm67394 Photorhab
53	44	54.3	463	4	ABG26615	Abg26615 Novel hum
54	44	54.3	489	8	ADL17928	Adl17928 Human sec
55	44	54.3	490	2	AAV13366	Aav13366 Amino aci
56	44	54.3	490	2	ADC78452	Adc78452 Human PRO
57	44	54.3	490	4	AAAB80234	Aaab80234 Human PRO
58	44	54.3	490	4	AAU12326	Aau12326 Human PRO
59	44	54.3	490	6	ABU71512	Abu71512 Human PRO
60	44	54.3	490	6	ABO17770	Ab017770 Novel hum
61	44	54.3	490	6	ABU71467	Abu71467 Human PRO
62	44	54.3	490	6	ABU81024	Abu81024 Human PRO
63	44	54.3	490	6	ABU71913	Abu71913 Human sec
64	44	54.3	490	6	ABO01796	Ab001796 Novel hum
65	44	54.3	490	6	ABU66724	Abu66724 Human PRO
66	44	54.3	490	6	ABU54369	Abu54369 Human sec
67	44	54.3	490	6	ABO47384	Ab047384 Human sec
68	44	54.3	490	6	ABU59805	Abu59805 Novel sec
69	44	54.3	490	6	ABO24995	Ab024995 Human sec
70	44	54.3	490	6	ABU64521	Abu64521 Human sec
71	44	54.3	490	6	ABU67367	Abu67367 Human sec
72	44	54.3	490	6	ABO14887	Ab014887 Human sec
73	44	54.3	490	6	ABU67000	Abu67000 Human sec
74	44	54.3	490	6	ABU69644	Abu69644 Novel hum
75	44	54.3	490	6	ABO14826	Ab014826 Human sec
76	44	54.3	490	6	ADA45829	Ada45829 Novel hum
77	44	54.3	490	6	ADA76260	Ada76260 Human PRO
78	44	54.3	490	6	ADB29337	Adb29337 Human sec
79	44	54.3	490	6	ADA18910	Ada18910 Human PRO
80	44	54.3	490	6	ADA61533	Ada61533 Homo sapi
81	44	54.3	490	6	ADB19318	Adb19318 Novel hum
82	44	54.3	490	6	ADB27859	Adb27859 Human PRO
83	44	54.3	490	6	ADA86338	Ada86338 Novel hum
84	44	54.3	490	6	ADA15902	Ada15902 Human PRO
85	44	54.3	490	6	ADA47688	Ada47688 Human PRO
86	44	54.3	490	6	ADA18193	Ada18193 Human sec
87	44	54.3	490	6	ABO32778	Ab032778 Human sec
88	44	54.3	490	6	ADA67483	Ada67483 Human PRO
89	44	54.3	490	6	ADA30490	Ada30490 Human PRO
90	44	54.3	490	6	ADA85786	Ada85786 Novel hum
91	44	54.3	490	6	ADA96998	Ada96998 Human PRO
92	44	54.3	490	6	ADA79302	Ada79302 Human PRO
93	44	54.3	490	6	ADA87441	Ada87441 Novel hum
94	44	54.3	490	6	ABO16643	Ab016643 Human PRO
95	44	54.3	490	6	ABO34838	Ab034838 Human PRO
96	44	54.3	490	6	ADA16168	Ada16168 Human sec
97	44	54.3	490	6	ADA91735	Ada91735 Novel hum

98	44	54.3	490	6	ADB14798	ADB14798 Human PRO	171	44	54.3	490	7	ADC34278	ADC34278 Human sec
99	44	54.3	490	6	ADB18759	ADB18759 Novel hum	172	44	54.3	490	7	ADC29333	ADC29333 Human sec
100	44	54.3	490	6	ADA93974	ADA93974 Human PRO	173	44	54.3	490	7	ADC28864	ADC28864 Human sec
101	44	54.3	490	6	ADB19870	ADB19870 Novel hum	174	44	54.3	490	7	ADC40749	ADC40749 Human sec
102	44	54.3	490	6	ADB13182	ADB13182 Human PRO	175	44	54.3	490	7	ADC19406	ADC19406 Human sec
103	44	54.3	490	6	ABO43303	ABO43303 Novel hum	176	44	54.3	490	7	ADC33854	ADC33854 Human sec
104	44	54.3	490	6	ADA74436	ADA74436 Human PRO	177	44	54.3	490	7	ADC12924	ADC12924 Human sec
105	44	54.3	490	6	ADA42313	ADA42313 Human sec	178	44	54.3	490	7	ADC50351	ADC50351 Novel hum
106	44	54.3	490	6	ADB24669	ADB24669 Human PRO	179	44	54.3	490	7	ADC71898	ADC71898 Novel hum
107	44	54.3	490	6	ADA82193	ADA82193 Human PRO	180	44	54.3	490	7	ADC59877	ADC59877 Novel hum
108	44	54.3	490	6	ADA75156	ADA75156 Human PRO	181	44	54.3	490	7	ADC52884	ADC52884 Novel hum
109	44	54.3	490	6	ADA85234	ADA85234 Novel hum	182	44	54.3	490	7	ADC57238	ADC57238 Novel hum
110	44	54.3	490	6	ADA84682	ADA84682 Novel hum	183	44	54.3	490	7	ADC60429	ADC60429 Novel hum
111	44	54.3	490	6	ABO17516	ABO17516 Human PRO	184	44	54.3	490	7	ADC50904	ADC50904 Novel hum
112	44	54.3	490	6	ADB29938	ADB29938 Human PRO	185	44	54.3	490	7	ADC65431	ADC65431 Human PRO
113	44	54.3	490	6	ADA80466	ADA80466 Human PRO	186	44	54.3	490	7	ADC54529	ADC54529 Novel hum
114	44	54.3	490	6	ADA75708	ADA75708 Human PRO	187	44	54.3	490	7	ADC53490	ADC53490 Novel hum
115	44	54.3	490	6	ADA46933	ADA46933 Human PRO	188	44	54.3	490	7	ADC59013	ADC59013 Novel hum
116	44	54.3	490	6	ADB25229	ADB25229 Human PRO	189	44	54.3	490	7	ADC55891	ADC55891 Novel hum
117	44	54.3	490	6	ADA93405	ADA93405 Human PRO	190	44	54.3	490	7	ADC58461	ADC58461 Novel hum
118	44	54.3	490	6	ADB26755	ADB26755 Human PRO	191	44	54.3	490	7	ADC12376	ADC12376 Human sec
119	44	54.3	490	6	ADB31042	ADB31042 Human PRO	192	44	54.3	490	7	ADD03135	ADD03135 Novel hum
120	44	54.3	490	6	ADA60970	ADA60970 Homo sapi	193	44	54.3	490	7	ADC90127	ADC90127 Novel hum
121	44	54.3	490	6	ADB24117	ADB24117 Human PRO	194	44	54.3	490	7	ADC69546	ADC69546 Human PRO
122	44	54.3	490	6	ADA96446	ADA96446 Human PRO	195	44	54.3	490	7	ADC48435	ADC48435 Human PRO
123	44	54.3	490	6	ADA81018	ADA81018 Human PRO	196	44	54.3	490	7	ADD09864	ADD09864 Human PRO
124	44	54.3	490	6	ADA95894	ADA95894 Human PRO	197	44	54.3	490	7	ADD04539	ADD04539 Novel hum
125	44	54.3	490	6	ADB26203	ADB26203 Human PRO	198	44	54.3	490	7	ADC80495	ADC80495 Novel hum
126	44	54.3	490	6	ADB21688	ADB21688 Novel hum	199	44	54.3	490	7	ADD11002	ADD11002 Human PRO
127	44	54.3	490	7	ADA77467	ADA77467 Human PRO	200	44	54.3	490	7	ADC47883	ADC47883 Human PRO
128	44	54.3	490	7	ADB18207	ADB18207 Human PRO	201	44	54.3	490	7	ADD04931	ADD04931 Human sec
129	44	54.3	490	7	ADA86890	ADA86890 Novel hum	202	44	54.3	490	7	ADC79943	ADC79943 Novel hum
130	44	54.3	490	7	ADA16592	ADA16592 Human sec	203	44	54.3	490	7	ADD09412	ADD09412 Human PRO
131	44	54.3	490	7	ADA13021	ADA13021 Human sec	204	44	54.3	490	7	ADD03937	ADD03937 Human sec
132	44	54.3	490	7	ADA41889	ADA41889 Human sec	205	44	54.3	490	7	ADD03513	ADD03513 Human sec
133	44	54.3	490	7	ADA87993	ADA87993 Novel hum	206	44	54.3	490	7	ADD41125	ADD41125 Novel hum
134	44	54.3	490	7	ADA46381	ADA46381 Novel hum	207	44	54.3	490	7	ADD52264	ADD52264 Human PRO
135	44	54.3	490	7	ADA17236	ADA17236 Human sec	208	44	54.3	490	7	ADD53004	ADD53004 Human PRO
136	44	54.3	490	7	ADA42739	ADA42739 Human sec	209	44	54.3	490	7	ADD53556	ADD53556 Novel hum
137	44	54.3	490	7	ADB28411	ADB28411 Human PRO	210	44	54.3	490	7	ADD51712	ADD51712 Human PRO
138	44	54.3	490	7	ADB28963	ADB28963 Human PRO	211	44	54.3	490	7	ADD02511	ADD02511 Human PRO
139	44	54.3	490	7	ADA76915	ADA76915 Human PRO	212	44	54.3	490	7	ADD01945	ADD01945 Human PRO
140	44	54.3	490	7	ADA88845	ADA88845 Novel hum	213	44	54.3	490	7	ADD54127	ADD54127 Novel hum
141	44	54.3	490	7	ADA97550	ADA97550 Human PRO	214	44	54.3	490	7	ADD92444	ADD92444 Human PRO
142	44	54.3	490	7	ADB27307	ADB27307 Human PRO	215	44	54.3	490	7	ADD91340	ADD91340 Human PRO
143	44	54.3	490	7	ADB22240	ADB22240 Novel hum	216	44	54.3	490	7	ADE03954	ADE03954 Human PRO
144	44	54.3	490	7	ABO17577	ABO17577 Human PRO	217	44	54.3	490	7	ADE32251	ADE32251 Novel hum
145	44	54.3	490	7	ADA66931	ADA66931 Human PRO	218	44	54.3	490	7	ADE22183	ADE22183 Human PRO
146	44	54.3	490	7	ADB22792	ADB22792 Human PRO	219	44	54.3	490	7	ADD79407	ADD79407 Human PRO
147	44	54.3	490	7	ADB23565	ADB23565 Human PRO	220	44	54.3	490	7	ADE41943	ADE41943 Human PRO
148	44	54.3	490	7	ADB15350	ADB15350 Human PRO	221	44	54.3	490	7	ADE17760	ADE17760 Human PRO
149	44	54.3	490	7	ADB15350	ADB15350 Human PRO	222	44	54.3	490	7	ADD91892	ADD91892 Human PRO
150	44	54.3	490	7	ADB38602	ADB38602 Novel hum	223	44	54.3	490	7	ADE33355	ADE33355 Novel hum
151	44	54.3	490	7	ADB38050	ADB38050 Novel hum	224	44	54.3	490	7	ADE333907	ADE333907 Novel hum
152	44	54.3	490	7	ADB66522	ADB66522 Novel hum	225	44	54.3	490	7	ADD79959	ADD79959 Human PRO
153	44	54.3	490	7	ADB89602	ADB89602 Human PRO	226	44	54.3	490	7	ADD92496	ADD92496 Human PRO
154	44	54.3	490	7	ADB90334	ADB90334 Human PRO	227	44	54.3	490	7	ADE19416	ADE19416 Human PRO
155	44	54.3	490	7	ADB77658	ADB77658 Human sec	228	44	54.3	490	7	ADE34765	ADE34765 Human sec
156	44	54.3	490	7	ADB39435	ADB39435 Novel hum	229	44	54.3	490	7	ADE18864	ADE18864 Human PRO
157	44	54.3	490	7	ADB74794	ADB74794 Human sec	230	44	54.3	490	7	ADE43060	ADE43060 Human PRO
158	44	54.3	490	7	ADB47058	ADB47058 Novel hum	231	44	54.3	490	7	ADD95849	ADD95849 Human PRO
159	44	54.3	490	7	ADB86665	ADB86665 Human PRO	232	44	54.3	490	7	ADE22735	ADE22735 Human PRO
160	44	54.3	490	7	ADB77270	ADB77270 Novel hum	233	44	54.3	490	7	ADD78853	ADD78853 Human PRO
161	44	54.3	490	7	ADB34427	ADB34427 Human PRO	234	44	54.3	490	7	ADE32803	ADE32803 Novel hum
162	44	54.3	490	7	ADB35531	ADB35531 Human PRO	235	44	54.3	490	7	ADE42495	ADE42495 Human PRO
163	44	54.3	490	7	ADB33875	ADB33875 Human PRO	236	44	54.3	490	7	ADD80511	ADD80511 Human PRO
164	44	54.3	490	7	ADB34979	ADB34979 Human PRO	237	44	54.3	490	7	ADD89539	ADD89539 Human PRO
165	44	54.3	490	7	ADB36083	ADB36083 Human PRO	238	44	54.3	490	7	ADD40823	ADD40823 Human PRO
166	44	54.3	490	7	ADB46478	ADB46478 Novel hum	239	44	54.3	490	7	ADE04622	ADE04622 Human PRO
167	44	54.3	490	7	ADB28440	ADB28440 Human sec	240	44	54.3	490	7	ADE92751	ADE92751 Human PRO
168	44	54.3	490	7	ADC39640	ADC39640 Human sec	241	44	54.3	490	7	ADG21460	ADG21460 Novel hum
169	44	54.3	490	7	ADC40154	ADC40154 Human sec	242	44	54.3	490	7	ADG23101	ADG23101 Novel hum
170	44	54.3	490	7	ADC18982	ADC18982 Human sec	243	44	54.3	490	7	ADF97436	ADF97436 Human PRO

244	44	54.3	490	7	ADG80500	Adg805000 Human PRO	317	44	54.3	490	8	ADG82156	Adg82156 Human PRO
245	44	54.3	490	7	ADG79948	Adg79948 Human PRO	318	44	54.3	490	8	ADG57395	Adg57395 Novel hum
246	44	54.3	490	7	ADH59248	Adh59248 Human sec	319	44	54.3	490	8	ADG56843	Adg56843 Novel hum
247	44	54.3	490	7	ADH55240	Adh55240 Novel hum	320	44	54.3	490	8	ADG55739	Adg55739 Novel hum
248	44	54.3	490	7	ADH55792	Adh55792 Novel hum	321	44	54.3	490	8	ADG58499	Adg58499 Novel hum
249	44	54.3	490	7	ADH55792	Adh55792 Human sec	322	44	54.3	490	8	ADG70865	Adg70865 Novel hum
250	44	54.3	490	7	ADI38027	Adi38027 Human sec	323	44	54.3	490	8	ADG92655	Adg92655 Human sec
251	44	54.3	490	7	ADI64011	Adi64011 Novel hum	324	44	54.3	490	8	ADG92655	Adg92655 Human sec
252	44	54.3	490	7	ADI63459	Adi63459 Novel hum	325	44	54.3	490	8	ADG57947	Adg57947 Novel hum
253	44	54.3	490	7	ADH81873	Adh81873 Novel hum	326	44	54.3	490	8	ADG53531	Adg53531 Novel hum
254	44	54.3	490	7	ADH81321	Adh81321 Novel hum	327	44	54.3	490	8	ADG711417	Adg711417 Novel hum
255	44	54.3	490	7	ADJ26295	Adj26295 Human sec	328	44	54.3	490	8	ADG81604	Adg81604 Human PRO
256	44	54.3	490	7	ADM04936	Adm04936 Human pro	329	44	54.3	490	8	ADH30566	Adh30566 Human PRO
257	44	54.3	490	7	ADM82490	Adm82490 Novel hum	330	44	54.3	490	8	ADH11933	Adh11933 Novel hum
258	44	54.3	490	7	ADN15689	Adn15689 Novel hum	331	44	54.3	490	8	ADG52355	Adg52355 Novel hum
259	44	54.3	490	7	ADN16518	Adn16518 Novel hum	332	44	54.3	490	8	ADG54083	Adg54083 Novel hum
260	44	54.3	490	7	ADN15337	Adn15337 Novel hum	333	44	54.3	490	8	ADG81052	Adg81052 Human PRO
261	44	54.3	490	7	ADN14785	Adn14785 Novel hum	334	44	54.3	490	8	ADG56291	Adg56291 Novel hum
262	44	54.3	490	7	ADN164960	Adn164960 Novel hum	335	44	54.3	490	8	ADH12557	Adh12557 Novel hum
263	44	54.3	490	7	ADC81047	Adc81047 Novel hum	336	44	54.3	490	8	ADH1403	Adh1403 Novel hum
264	44	54.3	490	7	ADG79210	Adg79210 Human sec	337	44	54.3	490	8	ADH28490	Adh28490 Human PRO
265	44	54.3	490	7	ADG76495	Adg76495 Human PRO	338	44	54.3	490	8	ADG54635	Adg54635 Novel hum
266	44	54.3	490	7	ADG7859	Adg7859 Human PRO	339	44	54.3	490	8	ADG59675	Adg59675 Novel hum
267	44	54.3	490	7	ADG86263	Adg86263 Human PRO	340	44	54.3	490	8	ADH20444	Adh20444 Human sec
268	44	54.3	490	7	ADG79634	Adg79634 Human PRO	341	44	54.3	490	8	ADH07299	Adh07299 Human sec
269	44	54.3	490	7	ADG75711	Adg75711 Human PRO	342	44	54.3	490	8	ADH59844	Adh59844 Human sec
270	44	54.3	490	7	ADG73310	Adg73310 Human PRO	343	44	54.3	490	8	ADH06872	Adh06872 Human sec
271	44	54.3	490	7	ADG23287	Adg23287 Human PRO	344	44	54.3	490	8	ADH181099	Adh181099 Human PRO
272	44	54.3	490	7	ADG23839	Adg23839 Human PRO	345	44	54.3	490	8	ADH16614	Adh16614 Human sec
273	44	54.3	490	7	ADG24482	Adg24482 Human PRO	346	44	54.3	490	8	ADH18614	Adh18614 Human sec
274	44	54.3	490	7	ADG87307	Adg87307 Human PRO	347	44	54.3	490	8	ADH165334	Adh165334 Human sec
275	44	54.3	490	7	ADG89173	Adg89173 Human PRO	348	44	54.3	490	8	ADH137597	Adh137597 Human sec
276	44	54.3	490	7	ADG73845	Adg73845 Human sec	349	44	54.3	490	8	ADG09842	Adg09842 Novel hum
277	44	54.3	490	7	ADG18312	Adg18312 Human PRO	350	44	54.3	490	8	ADH97393	Adh97393 Human sec
278	44	54.3	490	7	ADG88621	Adg88621 Human PRO	351	44	54.3	490	8	ADH15313	Adh15313 Novel hum
279	44	54.3	490	7	ADG99399	Adg99399 Human sec	352	44	54.3	490	8	ADG09190	Adg09190 Novel hum
280	44	54.3	490	7	ADG94641	Adg94641 Human PRO	353	44	54.3	490	8	ADH165761	Adh165761 Human sec
281	44	54.3	490	7	ADG91052	Adg91052 Human PRO	354	44	54.3	490	8	ADH14645	Adh14645 Novel hum
282	44	54.3	490	7	ADG95193	Adg95193 Human PRO	355	44	54.3	490	8	ADH18240	Adh18240 Novel hum
283	44	54.3	490	7	ADG93303	Adg93303 Human PRO	356	44	54.3	490	8	ADJ99561	Adj99561 Human sec
284	44	54.3	490	7	ADG34884	Adg34884 Human PRO	357	44	54.3	490	8	ADJ08754	Adj08754 Human sec
285	44	54.3	490	7	ADG98518	Adg98518 Human sec	358	44	54.3	490	8	ADJ25095	Adj25095 Human sec
286	44	54.3	490	7	ADG92199	Adg92199 Novel hum	359	44	54.3	490	8	ADJ63521	Adj63521 Novel hum
287	44	54.3	490	7	ADG90500	Adg90500 Human PRO	360	44	54.3	490	8	ADJ29845	Adj29845 Human sec
288	44	54.3	490	7	ADG91647	Adg91647 Novel hum	361	44	54.3	490	8	ADJ77416	Adj77416 Human PRO
289	44	54.3	490	7	ADG98945	Adg98945 Human sec	362	44	54.3	490	8	ADJ65538	Adj65538 Human PRO
290	44	54.3	490	7	ADG40415	Adg40415 Human sec	363	44	54.3	490	8	ADJ27674	Adj27674 Human PRO
291	44	54.3	490	7	ADG73809	Adg73809 Human sec	364	44	54.3	490	8	ADM27674	Adm27674 Human PRO
292	44	54.3	490	7	ADG22012	Adg22012 Novel hum	365	44	54.3	490	8	ADM42398	Adm42398 Human PRO
293	44	54.3	490	7	ADG20082	Adg20082 Human PRO	366	44	54.3	490	8	ADM42398	Adm42398 Human PRO
294	44	54.3	490	7	ADG97988	Adg97988 Human PRO	367	44	54.3	490	8	ADM42398	Adm42398 Human PRO
295	44	54.3	490	7	ADG24205	Adg24205 Novel hum	368	44	54.3	490	8	ADJ74567	Adj74567 Human sec
296	44	54.3	490	7	ADG33390	Adg33390 Human PRO	369	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
297	44	54.3	490	7	ADG98559	Adg98559 Human PRO	370	44	54.3	490	8	ADJ95742	Adj95742 Human sec
298	44	54.3	490	7	ADG99111	Adg99111 Human PRO	371	44	54.3	490	8	ADJ03604	Adj03604 Human sec
299	44	54.3	490	7	ADG16696	Adg16696 Human PRO	372	44	54.3	490	8	ADT03281	Adt03281 Human sec
300	44	54.3	490	7	ADG05155	Adg05155 Human PRO	373	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
301	44	54.3	490	7	ADG19422	Adg19422 Human PRO	374	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
302	44	54.3	490	7	ADG73385	Adg73385 Human sec	375	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
303	44	54.3	490	7	ADG13259	Adg13259 Human PRO	376	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
304	44	54.3	490	7	ADG08316	Adg08316 Novel hum	377	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
305	44	54.3	490	7	ADG15486	Adg15486 Human PRO	378	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
306	44	54.3	490	7	ADG96884	Adg96884 Human PRO	379	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
307	44	54.3	490	7	ADG06069	Adg06069 Human PRO	380	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
308	44	54.3	490	7	ADG23653	Adg23653 Novel hum	381	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
309	44	54.3	490	7	ADG03942	Adg03942 Human PRO	382	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
310	44	54.3	490	7	ADG24843	Adg24843 Novel hum	383	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
311	44	54.3	490	7	ADG07140	Adg07140 Novel hum	384	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
312	44	54.3	490	7	ADG07692	Adg07692 Novel hum	385	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
313	44	54.3	490	7	ADG55187	Adg55187 Novel hum	386	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
314	44	54.3	490	7	ADG60851	Adg60851 Novel hum	387	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
315	44	54.3	490	7	ADG61955	Adg61955 Novel hum	388	44	54.3	490	8	ADJ96294	Adj96294 Novel hum
316	44	54.3	490	7	ADG92228	Adg92228 Human sec	389	44	54.3	490	8	ADJ96294	Adj96294 Novel hum

390	42	51.9	145	8	ADJ32024	Adj32024 Plant ino	463	39	48.1	416	3	AAB53130	Aab53130 Macaca mu
391	42	51.9	159	3	AAV59431	Aav59431 Wheat ino	464	39	48.1	465	8	ADX77013	Adx77013 Plant ful
392	42	51.9	176	3	ADJ32015	Adj32015 Plant ino	465	39	48.1	484	9	Adx80686	Adx80686 Channel c
393	42	51.9	202	8	ADJ32026	Adj32026 Plant ino	466	39	48.1	533	6	ABU24808	Abu24808 Protein e
394	42	51.9	226	8	ADJ32019	Adj32019 Plant ino	467	39	48.1	591	4	ABG29309	Abg29309 Novel hum
395	42	51.9	244	3	AAV59422	AAV59422 Rice inos	468	39	48.1	765	6	ABU44799	Abu44799 Protein e
396	42	51.9	281	3	AAV59433	Aav59433 Wheat ino	469	39	48.1	791	8	ABM84773	Abm84773 Human dia
397	42	51.9	337	42	AAV59426	Aav59426 Soybean i	470	39	48.1	816	4	ABG25783	Abg25783 Novel hum
398	42	51.9	338	5	ABB92858	Abb92858 Herbicida	471	39	48.1	816	4	ABG04035	Abg04035 Novel hum
399	42	51.9	348	7	ADC13494	Adc13494 Protein w	472	39	48.1	831	4	RAM39264	Ram39264 Human pol
400	42	51.9	353	8	ADT56492	Adt56492 Plant pol	473	39	48.1	831	8	ADQ20117	Adq20117 Human sof
401	42	51.9	391	8	ADT55670	Adt55670 Plant pol	474	39	48.1	831	8	ADQ91460	Adq91460 Amino aci
402	42	51.9	403	8	ADX76324	Adx76324 Plant ful	475	39	48.1	831	9	ADY70286	Ady70286 Human bet
403	42	51.9	439	8	ADN20330	Adn20330 Bacterial	476	39	48.1	831	9	ADY70628	Ady70628 Human pae
404	42	51.9	789	7	ADM25991	Adm25991 Hyperther	477	39	48.1	833	8	ADO71748	Ado71748 BFLP1698
405	41.5	51.2	113	7	ADC00869	Adc00869 Enterohae	478	39	48.1	838	4	RAM41050	Ram41050 Human pol
406	41	50.6	14	2	AAW09542	Aaw09542 Thrombopo	479	39	48.1	840	8	ABM84774	Abm84774 Human dia
407	41	50.6	14	2	AAW36693	Aaw36693 Thrombopo	480	39	48.1	845	4	ABG21852	Abg21852 Novel hum
408	41	50.6	14	4	AAU25912	Aau25912 Human thr	481	39	48.1	880	8	ADL46155	Adl46155 Murine so
409	41	50.6	59	7	ABW73985	Abw73985 DNA clone	482	39	48.1	891	6	ABU49944	Abu49944 Protein e
410	41	50.6	82	5	ABP33075	Abp33075 Human ORF	483	39	48.1	892	4	AAU38250	Aau38250 Salmonell
411	41	50.6	100	8	ADS10566	Ads10566 Human che	484	39	48.1	892	6	ABU47096	Abu47096 Protein e
412	41	50.6	198	7	ADB80946	Adb80946 RING-SH c	485	39	48.1	892	6	ABU48057	Abu48057 Protein e
413	41	50.6	198	9	ADX07064	Adx07064 Cyclin-de	486	39	48.1	898	7	ABO63805	AbO63805 Klebsiell
414	41	50.6	240	8	ABM81851	Abm81851 Tumour-as	487	39	48.1	908	8	ADO71747	Ado71747 BFLP1698
415	41	50.6	470	2	AAW25766	Aaw25766 Human CAR	488	39	48.1	1019	8	ADO71746	Ado71746 BFLP1698
416	41	50.6	470	3	AAV98167	Aav98167 Human TRA	489	39	48.1	1019	8	ABM81255	Abm81255 Tumour-as
417	41	50.6	470	6	ABO07168	AbO07168 Human p53	490	39	48.1	1049	8	ADN18867	Adn18867 Bacterial
418	41	50.6	470	7	ADB80943	AdB80943 RING-SH c	491	39	48.1	1162	8	ADO71742	Ado71742 BFLP1698
419	41	50.6	470	8	ADP54388	Adp54388 Human PRO	492	39	48.1	1162	8	ADO71735	Ado71735 BFLP1698
420	41	50.6	470	9	ADX07062	Adx07062 Cyclin-de	493	39	48.1	1162	8	ADO71738	Ado71738 BFLP1698
421	41	50.6	470	9	ADY19800	Ady19800 PRO polyd	494	39	48.1	1162	8	ADO71744	Ado71744 BFLP1698
422	41	50.6	1885	8	ADL98343	Adl98343 SSL-1, SE	495	39	48.1	1162	8	ADO71737	Ado71737 BFLP1698
423	40	49.4	75	3	AAQ27112	Aaq27112 Zea mays	496	39	48.1	1162	8	ADO71740	Ado71740 BFLP1698
424	40	49.4	113	8	ADM87627	Adm87627 Human EST	497	39	48.1	1162	8	ADO71739	Ado71739 BFLP1698
425	40	49.4	143	3	AAQ16560	Aaq16560 Arabidops	498	39	48.1	1162	8	ADO71741	Ado71741 BFLP1698
426	40	49.4	185	9	ABP35749	Abp35749 L. pneumo	499	39	48.1	1162	8	ADO71743	Ado71743 BFLP1698
427	40	49.4	297	4	ABB12473	Abb12473 Human bon	500	39	48.1	1162	8	ADO71733	Ado71733 Human BFL
428	40	49.4	365	7	ADP58880	Adp58880 Human pol	501	39	48.1	1162	8	ADO71736	Ado71736 BFLP1698
429	40	49.4	374	7	ADC13502	Adc13502 Protein w	502	39	48.1	1162	8	ADO71745	Ado71745 BFLP1698
430	40	49.4	375	6	ABU31595	Abu31595 Protein e	503	39	48.1	1162	8	ADO71734	Ado71734 BFLP1698
431	40	49.4	384	8	ADY24810	Ady24810 Plant ful	504	39	48.1	1588	6	ABR52925	AbR52925 Protein s
432	40	49.4	391	7	ABO64064	AbO64064 Klebsiell	505	39	48.1	1588	7	ADK62502	AdK62502 Disease t
433	40	49.4	405	2	AAW70326	Aaw70326 Secreted	506	39	48.1	1588	8	ADN18982	Adn18982 Bacterial
434	40	49.4	427	9	ABM96829	Abm96829 M. xanthu	507	38.5	47.5	1858	5	ABB90859	Abb90859 Herbicida
435	40	49.4	461	6	ABR53263	AbR53263 Protein s	508	38	46.9	14	2	AAW09538	Aaw09538 Thrombopo
436	40	49.4	461	7	ADK63638	AdK63638 Disease t	509	38	46.9	14	2	AAW36689	Aaw36689 Thrombopo
437	40	49.4	461	8	ADS43875	AdS43875 Bacterial	510	38	46.9	116	7	ABO75047	AbO75047 Pseudomon
438	40	49.4	468	8	ADM87165	Adm87165 Human pro	511	38	46.9	143	3	ABG25206	Abg25206 Eucalyptu
439	40	49.4	497	6	ABP71435	Abp71435 B. halodu	512	38	46.9	225	4	AAQ98805	Aaq98805 Human cel
440	40	49.4	497	8	ADS28233	AdS28233 Bacterial	513	38	46.9	225	9	ABP36789	AbP36789 L. pneumo
441	40	49.4	521	8	ADI30117	Adi30117 M. methyl	514	38	46.9	225	9	ABE40164	AbE40164 L. pneumo
442	40	49.4	569	5	ABB94300	Abb94300 Chlamydia	515	38	46.9	240	4	RAM00973	Ram00973 Human bon
443	40	49.4	772	2	AAV51478	Aav51478 Human TLE	516	38	46.9	254	3	AAV83791	Aav83791 S. fradia
444	40	49.4	772	2	AAV51111	Aav51111 Human TLE	517	38	46.9	262	6	ABJ26540	AbJ26540 Aspergill
445	40	49.4	772	8	ADN04243	Adn04243 Antipsori	518	38	46.9	262	8	ADR59583	AdR59583 Aspergill
446	40	49.4	1044	6	ADA13361	Ada13361 Human int	519	38	46.9	277	1	AAQ70298	Aaq70298 Sequence
447	40	49.4	1093	8	ADO44174	AdO44174 Structura	520	38	46.9	277	2	AAW33275	Aaw33275 S. fradia
448	40	49.4	1101	4	AAW50136	Aaw50136 Human GTP	521	38	46.9	289	4	AAQ98803	Aaq98803 Human cel
449	40	49.4	1101	6	ADA13349	Ada13349 Human int	522	38	46.9	289	4	AAQ98800	Aaq98800 Human cel
450	39.5	48.8	113	7	ADC00042	Adc00042 Enterohae	523	38	46.9	311	4	AAQ98800	Aaq98800 Propionib
451	39.5	48.8	113	7	ADC00340	Adc00340 Enterohae	524	38	46.9	330	4	AAU48008	Aau48008 Propionib
452	39.5	48.8	113	7	ADC00493	Adc00493 Enterohae	525	38	46.9	330	6	ABM44527	Abm44527 Propionib
453	39	48.1	14	2	AAW09536	Aaw09536 Thrombopo	526	38	46.9	338	4	AAQ98797	Aaq98797 Human cel
454	39	48.1	14	2	AAW09544	Aaw09544 Thrombopo	527	38	46.9	375	4	AAQ98796	Aaq98796 Human cel
455	39	48.1	14	2	AAW36687	Aaw36687 Thrombopo	528	38	46.9	411	6	ABP79669	Abp79669 N. gonorr
456	39	48.1	14	2	AAW36695	Aaw36695 Thrombopo	529	38	46.9	419	4	AAQ98793	Aaq98793 Human cel
457	39	48.1	14	2	AAW36768	Aaw36768 Thrombopo	530	38	46.9	425	4	AAU35805	Aau35805 Helicobac
458	39	48.1	14	4	AAU25906	Aau25906 Human thr	531	38	46.9	425	6	ABU30955	Abu30955 Protein e
459	39	48.1	14	4	AAU25914	Aau25914 Human thr	532	38	46.9	447	6	ABU47669	Abu47669 Protein e
460	39	48.1	184	8	ADX67893	Adx67893 Plant ful	533	38	46.9	447	6	ABU44845	Abu44845 Protein e
461	39	48.1	218	2	AAW20717	Aaw20717 H. pylori	534	38	46.9	464	7	ADB70147	AdB70147 C. neofo
462	39	48.1	334	8	ADY12323	Ady12323 Plant ful	535	38	46.9	469	6	ABU40299	Abu40299 Protein e

536	38	46.9	469	6	ABU41961	Protein e	609	37	45.7	455	6	ABU23370	Protein e
537	38	46.9	471	6	ABU15562	Protein e	610	37	45.7	471	6	ABU21814	Protein e
538	38	46.9	471	6	ADU15046	Pseudomon	611	37	45.7	482	5	ADK34554	Novel hum
539	38	46.9	482	4	AAU27892	Human con	612	37	45.7	491	6	ABU22810	Protein e
540	38	46.9	488	4	AAG98787	Human cel	613	37	45.7	530	4	AAG90434	C glutami
541	38	46.9	490	4	AAG98786	Human cel	614	37	45.7	541	5	ABB57264	Mouse isc
542	38	46.9	522	8	ADT57198	Plant pol	615	37	45.7	542	7	ADD48911	Rat Prote
543	38	46.9	530	8	ADY07757	Plant ful	616	37	45.7	570	6	ABU62313	Rat MINT
544	38	46.9	555	7	ABO79567	Pseudomon	617	37	45.7	741	4	ABG19581	Novel hum
545	38	46.9	578	4	AAU27720	Human ful	618	37	45.7	749	4	ABG24612	Novel hum
546	38	46.9	604	6	ADA48262	Rice prot	619	37	45.7	750	6	ABU62312	Rat MINT
547	38	46.9	604	7	ABM86602	Rice abio	620	37	45.7	750	6	ABU62311	Rat Mx2
548	38	46.9	625	4	ABM95823	Human pro	621	37	45.7	766	4	ABG13668	Novel hum
549	38	46.9	764	2	AAW94359	Human Fac	622	37	45.7	773	5	AAE28503	Maize suc
550	38	46.9	893	6	ABU22898	Protein e	623	37	45.7	1026	5	ABP73754	Candida a
551	38	46.9	1243	8	ADR09787	Human pro	624	37	45.7	1027	2	AAW17057	Candida a
552	38	46.9	3262	3	AAV53675	Mechanica	625	37	45.7	1072	7	ABM85376	Mouse pro
553	38	46.9	3264	3	AAV53676	Protein 2	626	37	45.7	1115	5	ABW71230	Human leg
554	38	46.9	5183	6	AAE14793	Human mic	627	37	45.7	1115	7	ABW01535	Human leg
555	38	46.9	5183	6	ADO44006	Amino aci	628	37	45.7	1115	8	ADJ71905	Human lgs
556	38	46.9	5183	9	ADU06154	Cyclin-de	629	37	45.7	1244	8	ADN22906	Bacterial
557	37.5	46.3	12	2	AAW09524	Thrombopo	630	37	45.7	1244	8	ADN22907	Bacterial
558	37.5	46.3	12	2	AAW36675	Thrombopo	631	37	45.7	1244	8	ADN22908	Bacterial
559	37.5	46.3	12	3	ABU16993	TPO-mimet	632	37	45.7	1566	4	ABE64864	Human dia
560	37.5	46.3	12	4	AAU25894	Human thr	633	37	45.7	1577	8	ABM83607	Human dia
561	37.5	46.3	12	5	ABE72879	TPO mimet	634	37	45.7	1617	8	ABM83606	Human DIT
562	37.5	46.3	12	8	ADJ73030	TPO mimet	635	37	45.7	1653	6	ABR41799	Human MDD
563	37.5	46.3	12	8	ADJ52665	CHI delet	636	37	45.7	1653	6	ABU11522	Human MDD
564	37.5	46.3	12	8	ADJ51626	CHI delet	637	37	45.7	1682	4	ABG13667	Novel hum
565	37.5	46.3	16	2	AAW09464	Thrombopo	638	37	45.7	1682	4	ABG07239	Novel hum
566	37.5	46.3	16	2	AAW33329	Thrombopo	639	37	45.7	1911	2	AAK99534	Dermatomy
567	37.5	46.3	16	3	ABU17019	TPO-mimet	640	37	45.7	1912	4	AAK39288	Human pol
568	37.5	46.3	16	4	AAU25829	Human thr	641	37	45.7	1912	6	AAE36110	Human chr
569	37.5	46.3	16	5	ABE72905	TPO mimet	642	37	45.7	1912	6	AAE36109	Human chr
570	37.5	46.3	16	7	ADJ73057	TPO mimet	643	37	45.7	1912	8	ADP12552	Protein e
571	37.5	46.3	16	8	ADJ51653	CHI delet	644	37	45.7	1912	8	ABM81343	Tumour-as
572	37.5	46.3	221	7	ABO70412	Pseudomon	645	37	45.7	1936	4	ABG13664	Novel hum
573	37.5	46.3	262	6	ABU15700	Human met	646	37	45.7	2159	4	ABG13664	Novel hum
574	37.5	46.3	282	5	ADS98200	Protein f	647	37	45.7	2386	2	AAW13153	S. pombe
575	37.5	46.3	351	8	ADW09581	Human met	648	37	45.7	2386	8	ADN19726	Bacterial
576	37.5	46.3	351	8	ADW09581	Human met	649	37	45.7	2485	2	AAK59922	RAS assoc
577	37	45.7	12	2	AAW09582	Thrombopo	650	37	45.7	2485	2	AAK59921	RAS assoc
578	37	45.7	12	2	AAW36733	Thrombopo	651	37	45.7	2503	4	ABG13554	Novel hum
579	37	45.7	12	4	AAU25952	Thrombopo	652	37	45.7	2503	4	ABG07238	Novel hum
580	37	45.7	12	4	AAU25951	Human thr	653	37	45.7	2531	5	ADI16935	Rat NOVX
581	37	45.7	13	2	AAW36732	Thrombopo	654	37	45.7	2531	5	ADI16934	Rat NOVX
582	37	45.7	13	2	AAW36732	Thrombopo	655	37	45.7	2531	7	ADP63713	Rat Prote
583	37	45.7	57	4	AAW83901	Human sec	656	37	45.7	2531	7	ADP63705	Rat Prote
584	37	45.7	107	6	ABP75804	Human sec	657	37	45.7	2531	7	ADP63701	Rat Prote
585	37	45.7	126	3	ABG02882	Human sec	658	37	45.7	2531	7	ADP63701	Rat Prote
586	37	45.7	141	8	ADX74883	Plant ful	659	37	45.7	2818	2	AAW13280	Human neu
587	37	45.7	156	4	ABG19580	Novel hum	660	37	45.7	2818	2	AAW13280	Human neu
588	37	45.7	162	4	AAW90349	Human imm	661	37	45.7	2818	5	ABSO8078	Human neu
589	37	45.7	173	6	ABU23115	Protein e	662	37	45.7	2871	4	ABG13665	Novel hum
590	37	45.7	218	8	ADQ65547	Novel hum	663	37	45.7	2871	4	ABG13665	Novel hum
591	37	45.7	220	7	ABO80735	Pseudomon	664	37	45.7	2871	4	ABG07243	Novel hum
592	37	45.7	223	4	ABU11685	Human sec	665	37	45.7	2871	4	ABG13666	Novel hum
593	37	45.7	248	6	ABU40125	Phototrab	666	37	45.7	2891	4	ABG13666	Novel hum
594	37	45.7	273	6	ABM68436	Plant ful	667	37	45.7	2891	4	ABG07240	Novel hum
595	37	45.7	277	8	ADX78585	Human het	668	37	45.7	2891	4	ABG13556	Novel hum
596	37	45.7	291	5	ABW75700	Human pro	669	36.5	45.1	10	2	AAW09557	Thrombopo
597	37	45.7	291	7	ADM04528	Human pro	670	36.5	45.1	10	2	AAW36708	Thrombopo
598	37	45.7	304	6	ABU40110	Protein e	671	36.5	45.1	10	4	AAU25927	Human thr
599	37	45.7	308	5	ABU40110	Protein e	672	36.5	45.1	316	9	ABE27292	Human thr
600	37	45.7	312	4	ABG13680	Novel hum	673	36.5	45.1	348	4	ABE27292	Human nov
601	37	45.7	327	4	ABG16297	Novel hum	674	36.5	45.1	348	5	ABF51981	Human tra
602	37	45.7	333	3	AAV97409	Zebrafish	675	36.5	45.1	348	8	ADO32262	Novel hum
603	37	45.7	354	7	ABO78447	Pseudomon	676	36.5	45.1	350	2	AAW56265	Mouse rec
604	37	45.7	366	8	ADJ48863	Oil-associ	677	36.5	45.1	350	3	AAV91933	Murine Re
605	37	45.7	384	8	ADY11295	Plant ful	678	36.5	45.1	379	5	ABO92530	Herbicida
606	37	45.7	394	8	ADY11295	Plant ful	679	36.5	45.1	860	5	ABO7689	Rat gluta
607	37	45.7	412	6	ADA36849	Acinetoba	680	36	44.4	14	2	AAW09539	Thrombopo
608	37	45.7	417	2	AAW42401	TRAF2 bin	681	36	44.4	14	2	AAW36691	Thrombopo

682	36	44.4	14	2	AAW36690	Aaw36690	Thrombopo	755	36	44.4	335	4	AAB36650	Aab36650	Mouse cyt
683	36	44.4	14	2	AAW36769	Aaw36769	Thrombopo	756	36	44.4	335	4	AAE11196	Aae11196	Mouse int
684	36	44.4	14	3	AAB15996	Aab15996	TPO-mimet	757	36	44.4	335	4	AAE11194	Aae11194	Mouse int
685	36	44.4	14	4	AAU25909	Aau25909	Human chr	758	36	44.4	335	4	AAE11197	Aae11197	Mouse int
686	36	44.4	14	4	AAU25910	Aau25910	Human chr	759	36	44.4	335	6	ABU01016	Abu01016	Mouse cyt
687	36	44.4	14	5	ABB72882	Abb72882	TPO mimet	760	36	44.4	335	7	ADC07182	Adc07182	Mouse cyt
688	36	44.4	14	7	ADJ73033	Adj73033	TPO mimet	761	36	44.4	335	8	ADS00012	Ads00012	Murine in
689	36	44.4	14	8	ADJ52668	Adj52668	CH1 delet	762	36	44.4	335	9	ADW97245	Adw97245	Murine p4
690	36	44.4	14	8	ADJ51629	Adj51629	CH1 delet	763	36	44.4	335	9	ADZ65017	Adz65017	Mouse p40
691	36	44.4	85	4	AAU56464	Aau56464	Propionib	764	36	44.4	344	5	AAD77548	Aad77548	Human mem
692	36	44.4	85	6	AAU52983	Aau52983	Propionib	765	36	44.4	346	7	ADC31361	Adc31361	Human nov
693	36	44.4	99	3	AG01004	Ag01004	Human sec	766	36	44.4	350	8	ADT60598	Adt60598	Plant pol
694	36	44.4	140	7	ABO69232	Ab069232	Pseudomon	767	36	44.4	351	2	AAV35573	Aav35573	C. pneumo
695	36	44.4	157	4	AAW15531	Aaw15531	Peptide #	768	36	44.4	351	6	ABU27028	Abu27028	Protein e
696	36	44.4	157	4	ABB34538	Abb34538	Peptide #	769	36	44.4	353	8	ADX93161	Adx93161	Plant ful
697	36	44.4	157	4	AAW28021	Aaw28021	Peptide #	770	36	44.4	354	4	ABW72416	Abw72416	Murine DN
698	36	44.4	157	4	ABB29364	Abb29364	Peptide #	771	36	44.4	354	4	AAE13014	Aae13014	Mouse deo
699	36	44.4	157	4	ABB19947	Abb19947	Peptide #	772	36	44.4	354	4	ABG27268	Abg27268	Novel hum
700	36	44.4	157	4	AAW67721	Aaw67721	Human bon	773	36	44.4	357	7	ADB64961	Adb64961	Human pro
701	36	44.4	157	4	AAW55324	Aaw55324	Human bra	774	36	44.4	357	8	ADM39229	Adm39229	Human BNI
702	36	44.4	157	4	AG49359	Ag49359	Human liv	775	36	44.4	364	3	RAG48975	Rag48975	Arabidops
703	36	44.4	157	4	AAW03280	Aaw03280	Peptide #	776	36	44.4	364	5	ABB93221	Abb93221	Herbicida
704	36	44.4	159	2	AAW21745	Aaw21745	LECGNAB	777	36	44.4	369	8	ADQ65073	Adq65073	Novel hum
705	36	44.4	159	2	AAW83114	Aaw83114	Snowdrop	778	36	44.4	371	3	RAG48974	Rag48974	Arabidops
706	36	44.4	162	4	AAW73460	Aaw73460	Human gen	779	36	44.4	390	4	AAW66067	Aaw66067	Human TAN
707	36	44.4	162	5	ABB90156	Abb90156	Human pol	780	36	44.4	390	6	ABO32721	Ab032721	Secreted
708	36	44.4	183	4	AAW72613	Aaw72613	Human OR-	781	36	44.4	390	8	ADQ10397	Adq10397	Human pol
709	36	44.4	187	7	ABO73960	Ab073960	Pseudomon	782	36	44.4	390	8	ADQ98269	Adq98269	Human pol
710	36	44.4	198	8	ADX79170	Adx79170	Plant ful	783	36	44.4	397	3	AAV74889	Aav74889	Neisseria
711	36	44.4	203	6	ABP99502	Abp99502	Human sec	784	36	44.4	397	6	ABU37997	Abu37997	Protein e
712	36	44.4	203	6	ABR00975	AbR00975	Human gen	785	36	44.4	405	7	ADH08159	Adh08159	Novel pro
713	36	44.4	203	6	ADA98177	Ada98177	Human sec	786	36	44.4	405	7	ADH87580	Adh87580	Enterococ
714	36	44.4	203	6	ADA44031	Ada44031	Human sec	787	36	44.4	407	3	RAG48973	Rag48973	Arabidops
715	36	44.4	203	7	ADC20347	Adc20347	Human sec	788	36	44.4	409	4	AAE11931	Aae11931	Human CG1
716	36	44.4	203	7	ADF10687	Adf10687	Human sec	789	36	44.4	409	8	ADS10855	Ads10855	Human the
717	36	44.4	204	3	AAW51840	Aaw51840	Human sec	790	36	44.4	423	4	AAW66065	Aaw66065	Human TAN
718	36	44.4	205	3	AAW06132	Aaw06132	Caenorhab	791	36	44.4	423	6	ABU08369	Abu08369	Amino aci
719	36	44.4	221	4	AAW66068	Aaw66068	Human TAN	792	36	44.4	423	6	ABO32719	Ab032719	Secreted
720	36	44.4	221	6	ABO32722	Ab032722	Secreted	793	36	44.4	423	8	ADQ10395	Adq10395	Human pol
721	36	44.4	221	8	ADQ10398	Adq10398	Human pol	794	36	44.4	423	8	ADQ98267	Adq98267	Human pol
722	36	44.4	221	8	ADQ98270	Adq98270	Human pol	795	36	44.4	423	8	ADS10856	Ads10856	Human the
723	36	44.4	230	7	ADB69992	Adb69992	C. neofo	796	36	44.4	458	8	ADQ09164	Adq09164	Novel pro
724	36	44.4	239	8	ADY22640	Ady22640	Plant ful	797	36	44.4	465	7	ADS12105	Ads12105	Human the
725	36	44.4	250	8	ADG22312	Adg22312	Cyanophag	798	36	44.4	465	8	ADS12106	Ads12106	Human the
726	36	44.4	255	7	ADC33028	Adc33028	Human nov	799	36	44.4	466	8	ADH22527	Adh22527	Human tra
727	36	44.4	275	5	ABP56265	Abp56265	Human NIP	800	36	44.4	468	4	ABB70033	Abb70033	Drosophil
728	36	44.4	275	8	ADM39231	Adm39231	Human BNI	801	36	44.4	479	9	ADW17429	Adw17429	Eucalyptu
729	36	44.4	282	2	AAW11871	Aaw11871	Recombina	802	36	44.4	480	4	ABB66251	Abb66251	Drosophil
730	36	44.4	282	8	ABM81515	Abm81515	Tumour-as	803	36	44.4	480	4	ABB70230	Abb70230	Drosophil
731	36	44.4	282	9	AAW08788	Aaw08788	Human pro	804	36	44.4	497	5	AAW49570	Aaw49570	B. napue
732	36	44.4	282	9	AAW17049	Aaw17049	Alzheimer	805	36	44.4	497	7	ADC25951	Adc25951	Oilseed r
733	36	44.4	287	2	AAW51096	Aaw51096	Ehrlichia	806	36	44.4	513	8	ADN20604	Adn20604	Bacterial
734	36	44.4	287	3	AAW06960	Aaw06960	E. canis	807	36	44.4	513	8	ADB10524	Adb10524	Allioococ
735	36	44.4	287	3	AAW71481	Aaw71481	Ehrlichia	808	36	44.4	531	8	ADQ14471	Adq14471	Mouse hyp
736	36	44.4	287	3	AAW36190	Aaw36190	Ehrlichia	809	36	44.4	531	8	ADS73606	Ads73606	Mouse IL-
737	36	44.4	287	4	AAU04200	Aau04200	Variable	810	36	44.4	538	7	ABO73681	Ab073681	Pseudomon
738	36	44.4	287	5	AAU96104	Aau96104	Ehrlichia	811	36	44.4	541	6	ABU89759	Abu89759	Protein d
739	36	44.4	287	5	ABG77951	Abg77951	Ehrlichia	812	36	44.4	541	6	ADU48913	Adu48913	Human Pro
740	36	44.4	287	6	ADA09767	Ada09767	E. canis	813	36	44.4	541	8	ADR09709	Adr09709	Human pro
741	36	44.4	287	9	ADW04260	Adw04260	Ehrlichia	814	36	44.4	555	6	ABB99660	Abb99660	Amino aci
742	36	44.4	295	7	ADM27021	Adm27021	Hyperther	815	36	44.4	569	5	ABP40727	Abp40727	Staphyloc
743	36	44.4	298	8	ADY09054	Ady09054	Plant ful	816	36	44.4	569	8	ADS08025	Ads08025	Staphyloc
744	36	44.4	302	6	ABP71305	Abp71305	Murine on	817	36	44.4	572	8	ADG31207	Adg31207	Novel nou
745	36	44.4	303	6	ABU37060	Abu37060	Human bre	818	36	44.4	572	8	ADQ67668	Adq67668	Novel hum
746	36	44.4	303	6	ABP71304	Abp71304	Human onc	819	36	44.4	573	4	AAW97498	Aaw97498	Murine DN
747	36	44.4	303	8	ADQ18331	Adq18331	Human sof	820	36	44.4	595	6	ABU22647	Abu22647	Protein e
748	36	44.4	307	8	ADR14487	Adr14487	Human NF-	821	36	44.4	595	6	ABU19799	Abu19799	Protein e
749	36	44.4	317	7	ADC97305	Adc97305	E. faeciu	822	36	44.4	595	6	ABU21867	Abu21867	Protein e
750	36	44.4	320	3	AAW10428	Aaw10428	Arabidops	823	36	44.4	597	3	RAG30765	Rag30765	Arabidops
751	36	44.4	322	7	ADF04707	Adf04707	Bacterial	824	36	44.4	617	3	ADC31032	Adc31032	Human nov
752	36	44.4	335	2	AAW49247	Aaw49247	Mouse int	825	36	44.4	619	8	ADQ67755	Adq67755	Novel hum
753	36	44.4	335	2	AAW44005	Aaw44005	Murine in	826	36	44.4	621	4	AAW94049	Aaw94049	Human pro
754	36	44.4	335	2	AAW56710	Aaw56710	p35 subun	827	36	44.4	638	7	ADK40958	Adk40958	Novel hum

974 35 43.2 315 8 ADP23634
 975 35 43.2 315 8 ADS15841
 976 35 43.2 315 9 ADY19896
 977 35 43.2 315 9 ADY15230
 978 35 43.2 319 4 AAB76547
 979 35 43.2 321 4 AAB76546
 980 35 43.2 321 4 AAG91586
 981 35 43.2 321 7 ADL65571
 982 35 43.2 321 8 ADS41652
 983 35 43.2 322 2 AAR47859
 984 35 43.2 322 2 AAY41120
 985 35 43.2 322 2 AAY22599
 986 35 43.2 329 6 ABU23480
 987 35 43.2 335 2 AAW80661
 988 35 43.2 339 4 ABB58379
 989 35 43.2 339 8 ADX88353
 990 35 43.2 343 4 AAG72916
 991 35 43.2 346 4 AAU58400
 992 35 43.2 346 6 ABM54919
 993 35 43.2 351 8 ADM41707
 994 35 43.2 356 2 AAR48547
 995 35 43.2 357 6 ABM68597
 996 35 43.2 363 7 ADJ68958
 997 35 43.2 367 8 ADP04625
 998 35 43.2 377 6 ABU41138
 999 35 43.2 385 7 ADF07153
 1000 35 43.2 394 8 ADX67615

ALIGNMENTS

RESULT 1
 AAW09535
 ID AAW09535 standard; protein; 14 AA.

AC AAW09535;

DT 10-SEP-1997 (first entry)

DE Thrombopoietin receptor binding peptide.

KW Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.

OS Synthetic.

PN WO9640189-A1.

PD 19-DEC-1996.

PF 05-JUN-1996; 96WO-US008998.

PR 07-JUN-1995; 95US-00472371.

PR 07-JUN-1995; 95US-00473604.

PR 07-JUN-1995; 95US-00476168.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00484090.

PA (GLAX) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PT Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

PS WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.
 XX thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 28; 106pp; English.

ADP23634 PRO polyp
 ADS15841 Human spr
 ADY19896 PRO polyp
 ADY15230 PRO polyp
 AAB76547 Coryneb
 AAB76546 Coryneb
 AAG91586 C glutami
 ADL65571 C. glutam
 ADS41652 Bacterial
 AAR47859 Human LDL
 AAY41120 Soybean s
 AAY22599 LDL recep
 ABU23480 Protein e
 AAW80661 S. pneumo
 ABB58379 Drosophil
 ADX88353 Plant ful
 AAG72916 Human olf
 AAU58400 Propionib
 ABM54919 Propionib
 ADM41707 Protein m
 AAR48547 Sequence
 ABM68597 Photorhab
 ADJ68958 Human hea
 ADP04625 Sea squir
 ABU41138 Protein e
 ADF07153 Bacterial
 ADX67615 Plant ful

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells

XX Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFPLSHGGYVC 14

DB 1 CSLQEFPLSHGGYVC 14

RESULT 2

AAW36686
 ID AAW36686 standard; peptide; 14 AA.

XX AAW36686;

DT 27-FEB-1998 (first entry)

DE Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

OS Synthetic.

PN WO9640750-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US009623.

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PA (GLAX) GLAXO GROUP LTD.

PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

PT Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

PS WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

PS Disclosure; Page 28; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines

XX Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
|||||
Db 1 CSLQEFSLSHGGYVC 14

RESULT 3
AAB16997
ID AAB16997 standard; peptide; 14 AA.
XX
AC AAB16997;
XX

31-OCT-2000 (first entry)

TPO-mimetic peptide sequence SEQ ID NO:53.

Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
immunosuppressive; BPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
inhibitor; erythropoietin; thrombopoietin; interleukin 1;
cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
vascular endothelial growth factor; matrix metalloproteinase; asthma;
thrombosis; pharmaceutical.

Synthetic.

WO200024782-A2.

04-MAY-2000.

25-OCT-1999; 99WO-US025044.

23-OCT-1998; 98US-0105371P.

22-OCT-1999; 99US-00428082.

(AMGE-) AMGEN INC.

Feige U, Liu C, Cheetham J, Boone TC;

WPI; 2000-350702/30.

Novel composition of matter comprising an Fc domain and pharmacologically active peptides, useful for treating cancer and autoimmune diseases.

Claim 19; Page 213; 608pp; English.

The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a, e, and f is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAG69443 to AAG69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

Sequence 14 AA;

Query Match 100.0%; Score 81; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
|||||
Db 1 CSLQEFSLSHGGYVC 14

RESULT 4

AAU25905

ID AAU25905 standard; peptide; 14 AA.

XX

AC AAU25905;

XX

DT 17-DEC-2001 (first entry)

XX

DE Human thrombopoietin receptor (TPO-R) activator peptide #91.

XX

Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine; haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA; bone marrow transplantation; haematological disorder; platelet disorder; enzyme-linked immunosorbent assay; in situ staining; biological fluid; tissue homogenate; fluorescence-activated cell sorting; Western blotting; in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

Homo sapiens.

OS

XX

PN US6251864-B1.

XX

PD 26-JUN-2001.

XX

PF 01-MAR-2000; 2000US-00516704.

XX

PR 07-JUN-1995; 95US-00478128.

PR 07-JUN-1995; 95US-00485301.

PR 07-JUN-1996; 96WO-US009623.

PR 15-AUG-1996; 96US-00699027.

XX

PA (GLAXO) GLAXO GROUP LTD.

XX

PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;

PI Balasubramanian P, Wagstrom CR, Hendren RW, Poddaturi S;

PI Yin Q;

XX

DR WPI; 2001-564142/63.

XX

Activating thrombopoietin receptors in cells, used to treat thrombocytopenia and hematological disorders, comprises contacting cells with peptides and peptide mimetics attached to hydrophilic polymers.

Disclosure; Col 21; 128pp; English.

Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that bind to and activate the human thrombopoietin receptor (TPO-R). Methods of activating thrombopoietin receptors in cells comprise contacting the cells with effective amounts of peptides and peptide mimetics attached to hydrophilic polymers. The methods are used to treat thrombocytopenia such as that due to chemotherapy, radiation therapy or bone-marrow transplantation and to prevent thrombocytopenia in patients at risk. The sequences are used to treat and prevent haematological disorders including thrombocytopenia and platelet disorders. They are used in vitro as unique tools for understanding the biological role of thrombopoietin (TPO) and to develop other compounds that bind to and activate the TPO receptor. The peptides can be used to detect TPO receptors on living cells and fixed cells, in biological fluids, in tissue homogenates, and in purified or natural biological materials. They may also be used for in situ staining, fluorescence-activated cell sorting, Western blotting and enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can be used for in vitro expansion of megakaryocytes and their committed progenitors alone or in conjunction with additional cytokines

Sequence 14 AA;

Query Match 100.0%; Score 81; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
|||||

Db 1 CSLQFLSHGGYVC 14

RESULT 5

ABB72883

ID ABB72883 standard; peptide; 14 AA.

AC ABB72883;

XX

DT 05-APR-2002 (first entry)

XX

DE TPO mimetic peptide SEQ ID NO:53.

XX

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; Igg; EPO;

KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;

KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;

KW TPO mimetic peptide; EPO mimetic peptide; BMP; VEGF antagonist;

KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;

KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;

KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;

KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;

KW sleep disorder; neurological degenerative disease; anaemia;

KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;

KW Fanconi's syndrome.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200183525-A2.

XX

PD 08-NOV-2001.

XX

PP 02-MAY-2001; 2001WO-US014310.

XX

PR 03-MAY-2000; 2000US-00563286.

XX

PA (AMGE-) AMGEN INC.

XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX

DR WPI; 2002-130313/17.

XX

PT Novel vehicle-peptide molecule or its multimers useful for treating

PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,

PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX

PS Claim 39; Page 43; 176pp; English.

XX

CC The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,

CC cycostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,

CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and

CC neuroprotective activities. (I) can be used as a therapeutic or

CC prophylactic agent as well as for screening purposes. (I) is useful for

CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of

CC interest, as a part of diagnostic kit to detect the presence of their

CC proteins of interest in a biological sample. Additionally, (I) is useful

CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,

CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,

CC infertility, and neurological degenerative diseases. (I), comprising EPO-

CC mimetic compounds are useful for treating disorders characterised by low

CC red blood cell levels such as anaemia. The TPO-mimetic comprising

CC compounds are useful for treating conditions that involve an existing

CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet

CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the

CC exemplification of the present invention

XX

Sequence 14 AA;

Query Match 100.0%; Score 81; DB 7; Length 14;

Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSLQFLSHGGYVC 14

Db 1 CSLQFLSHGGYVC 14

RESULT 6

ADJ73034

ID ADJ73034 standard; peptide; 14 AA.

XX

AC ADJ73034;

XX

DT 06-MAY-2004 (first entry)

XX

DE TPO mimetic peptide sequence SeqID 488.

XX

KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;

KW cardiovascular; infectious; malignant; neurologic disease; anaemia;

KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;

KW TPO.

XX

OS Synthetic.

XX

PN WO2003084477-A2.

XX

PD 16-OCT-2003.

XX

PF 24-MAR-2003; 2003WO-US009139.

XX

PR 29-MAR-2002; 2002US-0368791P.

XX

PA (CENZ) CENTOCOR INC.

XX

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX

DR WPI; 2003-804237/75.

XX

PT New CDR mimetibody comprising a portion of a heavy or light chain

PT variable region comprising human framework or ligand binding region,

PT useful for preparing a composition for treating e.g., immune,

PT cardiovascular or neurologic disease.

XX

PS Disclosure; SEQ ID NO 488; 97pp; English.

XX

CC This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion

CC of, one or more CDR regions, such that each CDR mimetibody comprises at

CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one

CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products

CC that can be useful in gene therapy and the generation of transgenic

CC plants and animals. Furthermore, the CDR mimetibody is useful for

CC preparing compositions for modulating, treating or reducing the symptoms

CC of immune, cardiovascular, infectious, malignant and/ or neurologic

CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,

CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This

CC peptide sequence is a TPO mimetic peptide sequence used to make a

CC mimetibody of the invention.

XX

Sequence 14 AA;

Query Match 100.0%; Score 81; DB 7; Length 14;

Best Local Similarity 100.0%; Pred. No. 2.8e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSLQFLSHGGYVC 14

Db 1 CSLQFLSHGGYVC 14

Db	1	CSLQEFSLHGGYVC 14	
RESULT 7			
ADJ52669			
ID	ADJ52669	standard; peptide; 14 AA.	
XX	AC	ADJ52669;	
XX	DT	06-MAY-2004 (first entry)	
XX	DE	CH1 deleted mimetibody-related peptide SeqID488.	
XX	KW	CH1 deleted mimetibody; immunosuppressive; cardiovascular; candida;	
XX	KW	hypotensive; neuroprotective; nontropic; antibacterial; virucide;	
XX	KW	fungicide; gene therapy; immune disorder; cardiovascular disease;	
XX	KW	arrhythmia; hypertension; heart failure; neurodegenerative;	
XX	KW	multiple sclerosis; dementia; Alzheimer's disease; anaemia;	
XX	KW	cancerous condition; infectious disease; bacterial infection;	
XX	KW	viral infection; fungal infection.	
XX	OS	Unidentified.	
XX	OS	Synthetic.	
XX	PN	WO2004002417-A2.	
XX	PD	08-JAN-2004.	
XX	PF	27-JUN-2003; 2003WO-US020347.	
XX	PR	28-JUN-2002; 2002US-0392431P.	
XX	PA	(CENZ) CENTOCOR INC.	
XX	PI	Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neseppor TC;	
XX	PI	Kutoloski KA;	
XX	DR	WPI; 2004-082870/08.	
XX	XX	New CH1-deleted mimetibody polypeptides and nucleic acids, useful for	
XX	PT	modulating, treating, alleviating, preventing an immune, cardiovascular,	
XX	PT	or neurodegenerative disease or disorder, anemia, cancer, or infectious	
XX	PT	diseases.	
XX	PS	Claim 2; SEQ ID NO 488; 129pp; English.	
XX	CC	This invention relates to CH1 deleted mimetibodies (and the DNA sequences	
XX	CC	which encode them), compositions, methods and uses. The invention may be	
XX	CC	useful for the development of compounds with an immunosuppressive,	
XX	CC	cardiovascular, cardiant, hypotensive, neuroprotective, nontropic,	
XX	CC	antibacterial, virucide or fungicide activity. In addition, the disclosed	
XX	CC	sequences may prove useful for gene therapy. The CH1-deleted mimetibody	
XX	CC	is useful for diagnosing or treating a disease condition in a cell,	
XX	CC	tissue, organ or animal, specifically for modulating, treating,	
XX	CC	alleviating, preventing the incidence or reducing the symptoms, of an	
XX	CC	immune, cardiovascular (for example arrhythmia, hypertension or heart	
XX	CC	failure), or neurodegenerative (for example multiple sclerosis, dementia	
XX	CC	or Alzheimer's disease) diseases or disorders, anaemia, cancerous	
XX	CC	conditions, or infectious diseases (for example bacterial, viral or	
XX	CC	fungal infection). The present sequence is that of a peptide which may be	
XX	CC	used during the creation of a mimetibody of the invention.	
XX	SQ	Sequence 14 AA;	
		Query Match 100.0%; Score 81; DB 8; Length 14;	
		Best Local Similarity 100.0%; Pred. No. 2.8e-06;	
		Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	CSLQEFSLHGGYVC 14	
Db	1	CSLQEFSLHGGYVC 14	
RESULT 8			
ADJ51630			
ID	ADJ51630	standard; peptide; 14 AA.	
XX	AC	ADJ51630;	
XX	DT	06-MAY-2004 (first entry)	
XX	DE	CH1 deleted mimetibody-related peptide SeqID488.	
XX	XX	CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;	
XX	KW	dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;	
XX	KW	gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;	
XX	KW	antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;	
XX	KW	ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;	
XX	KW	TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;	
XX	KW	dental disorder; oral disorder; dermatological disorder; ear disorder;	
XX	KW	nose disorder; throat disorder; endocrine disorder; metabolic disorder;	
XX	KW	gastrointestinal disorder; gynaecological disorder; hepatic disorder;	
XX	KW	obstetric disorder; haematologic disorder; musculoskeletal disorder;	
XX	KW	allergic disorder; infectious disorder; neurological disorder;	
XX	KW	oncological disorder; neurological disorder; nutritional disorder;	
XX	KW	ophthalmologic disorder; pediatric disorder; psychiatric disorder;	
XX	KW	renal disorder; pulmonary disorder.	
XX	OS	Unidentified.	
XX	OS	Synthetic.	
XX	PN	WO2004002424-A2.	
XX	PD	08-JAN-2004.	
XX	XX	30-JUN-2003; 2003WO-US020495.	
XX	PF	28-JUN-2002; 2002US-0392431P.	
XX	PR	19-SEP-2002; 2002US-0412144P.	
XX	XX	(CENZ) CENTOCOR INC.	
XX	XX	Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neseppor TC;	
XX	XX	Kutoloski KA;	
XX	DR	WPI; 2004-082872/08.	
XX	XX	New CH1 deleted mimetibody polypeptide and nucleic acid, useful for	
XX	PT	diagnosing, preventing or treating cardiovascular, dermatologic,	
XX	PT	endocrine, gastrointestinal, gynecologic, infectious, neurologic and	
XX	PT	nutritional disorders.	
XX	PS	Claim 14; SEQ ID NO 488; 123pp; English.	
XX	CC	This invention relates to CH1 deleted mimetibodies (and the DNA sequences	
XX	CC	which encode them), compositions, methods and uses. The invention may be	
XX	CC	useful for the development of compounds with an osteopathic,	
XX	CC	cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,	
XX	CC	gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,	
XX	CC	immunomodulator, antiallergic, muscular-Gen, cytostatic,	
XX	CC	antiinflammatory, neuroleptic, ophthalmological, nephrotropic or	
XX	CC	respiratory-Gen activity acting as a tumour necrosis factor (TNF)-	
XX	CC	modulator or cytokine-agonist. The methods and compositions of the	
XX	CC	present invention are useful for the diagnosis, prevention and/or	
XX	CC	treatment of diseases or conditions associated with aberrant expression	
XX	CC	or activity of the CH1 deleted mimetibody, such as a bone or joint,	
XX	CC	cardiovascular, dental or oral, dermatological, ear, nose or throat,	
XX	CC	endocrine, metabolic, gastrointestinal, gynaecological, hepatic,	
XX	CC	obstetric, haematologic, immunological, allergic, infectious,	
XX	CC	musculoskeletal, oncological, neurological, nutritional, ophthalmologic,	
XX	CC	pediatric, psychiatric, renal or pulmonary disorders. The present	
XX	CC	sequence is that of a peptide which may be used during the creation of a	
XX	CC	mimetibody of the invention.	
XX	SQ	Sequence 14 AA;	

Query Match 100.0%; Score 81; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGQYVC 14
DB 1 CSLQEFLSHGQYVC 14

RESULT 9
ADT60670
ID ADT60670 standard; protein; 453 AA.
XX
XX ADT60670;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant polypeptide, SEQ ID 10747.
XX
XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
XX Viridiplantae.
XX
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
XX Claim 2; SEQ ID NO 10747; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of

the present invention are useful in the field of biochemistry and
genetics, in particular for producing transgenic plants with improved
biological characteristics such as increased yield, improved nitrogen
flow, increasing plant tolerance to cold or heat, improving plant
tolerance to extreme osmotic and drought conditions, and improving plant
tolerance to plant pests or pathogens. They can also be used in physical
arrays of molecules, plant breeding markers, computer-based storage and
analysis systems. The present sequence is one of the 5544 plant protein
sequences of the invention. Note: The sequence data for this patent did
not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 453 AA;

Query Match 60.5%; Score 49; DB 8; Length 453;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLQEFLSHGQYVC 14
DB 140 CALKWFISHGSTVC 153

RESULT 10
AAY96482
ID AAY96482 standard; protein; 344 AA.
XX
XX AAY96482;
XX
DT 12-SEP-2000 (first entry)
XX
XX Human topoisomerase III-alpha interacting protein.
DE
XX Topoisomerase III-alpha; interacting protein; TIII-a IP; RNA helicase;
KW DBX1; nucleosome; supercoiling; chromosome segregation; recombination;
KW stabilization; cell division; apoptosis; cell cycle regulation;
KW cytotoxic; anti-tumour.
XX
XX Homo sapiens.
OS
XX WO200032768-A1.
XX
XX 08-JUN-2000.
XX
XX 29-NOV-1999; 99WO-FR002952.
PF
XX 30-NOV-1998; 98FR-00015081.
PR
XX (AVET) AVENTIS PHARMA SA.
XX
XX Fournier A, Goulaouic H, Riou J;
PI WPI; 2000-412316/35.
XX
XX N-PSDB; AAA29206.
DR
XX New nucleic acid encoding ligand for topoisomerase III alpha, useful for
PT inhibiting the enzyme and in drug screening, e.g. for potential
PT anticancer agents.
XX
XX Claim 5; Page 46-47; 68pp; French.
XX
XX This shows the human topoisomerase III-alpha interacting protein (TIII-a
CC IPI). TIII-a IP has structural features in common with RNA helicases (e.g.
CC DBX1), which are involved in destruction of the nucleosome structure,
CC supercoiling of DNA, segregation of newly replicated chromosomes, and
CC recombination and stabilization of the genome. Agents that alter
CC interaction between TIII-a IP and TIII-a therefore modulate cell
CC division, replication, transcription, translation, splicing and DNA
CC recombination or repair, so may slow cell growth, block the cell cycle or
CC induce apoptosis. Antibodies and ligands of TIII-a IP are used to
CC prevent, treat or alleviate diseases that involve abnormal regulation of
CC the cell cycle, i.e. they are potential anti-tumour (cytotoxic) agents

```
XX SQ Sequence 344 AA;
Query Match 59.3%; Score 48; DB 3; Length 344;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
Db 139 SLEDFLYHEGYAC 151

RESULT 11
AAW81503
ID AAW81503 standard; protein; 660 AA.
XX
AC AAW81503;
DT 02-FEB-1999 (first entry)
XX
DE Dead Box Y (DBY) gene product.
XX
KW Non-recombining region; human; Y chromosome; X homologue; testis; DBY;
KW infertility; sperm; gene alteration; inhibitor; Dead Box Y.
XX
OS Homo sapiens.
XX
PN WO9846747-A2.
XX
PD 22-OCT-1998.
XX
PF 10-APR-1998; 98WO-US007115.
XX
PR 11-APR-1997; 97US-0041877P.
XX
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Lahn BT, Page DC;
XX
DR WPI; 1998-568729/48.
XX
DR N-PSDB; AAV69633.
XX
Novel genes in the non-combining region of Y chromosome - useful to
diagnose if male infertility or reduced sperm count has a genetic basis.
XX
PS Claim 6; Fig 3A-B; 54pp; English.
XX
This represents the amino acid sequence of the Dead Box Y (DBY) gene
product. The invention relates to genes occurring on the non-recombining
region of the human Y chromosome. The sequences fall into two classes:
(1) X-homologous DNA which are expressed in many organs, having
functional X homologues and (2) testis-specific DNA sequences. Y
chromosomal DNA from males with known conditions such as infertility and
reduced sperm count can be assessed using the invention to determine
whether the condition is associated with or caused by the occurrence of
the gene or gene alteration. Candidate inhibitors of the enzymatic
activity of the genes can be assessed using in vitro assays
XX
SQ Sequence 660 AA;
Query Match 59.3%; Score 48; DB 2; Length 660;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
Db 454 SLEDFLYHEGYAC 466

RESULT 12
ABO07237
ID ABO07237 standard; protein; 660 AA.
XX
```

```
AC ABO07237;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 197.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
DR WPI; 2003-156859/15.
DR N-PSDB; ACD13410.
XX
Identifying modulators of the p53 pathway for use in treating apoptotic
or cell proliferation disorders, comprises screening for agents that
modulate activity of a human ortholog of genes that modify the p53
pathway in Drosophila.
XX
Example 2; Page 567-569; 678pp; English.
XX
The invention relates to identifying (M1) a candidate p53 pathway
modulating agent, by contacting an assay system comprising a purified HM
polypeptide (human orthologue of genes that modify the p53 pathway in
Drosophila) or nucleic acid with a test agent under conditions, where but
for the presence of the test agent, the system provides a reference
activity, and detecting a test agent-biased activity of a cell (comprising
Also included are modulating (M2) a p53 pathway of a cell (comprising
contacting a cell defective in p53 function with a candidate modulator
that specifically binds to a HM polypeptide comprising an HM amino acid
sequence, where p53 function is restored), modulating (M3) a p53 pathway
in a mammalian cell (comprising contacting the cell with an agent that
specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
a disease in a patient (comprising: (a) obtaining a biological sample
from the patient; (b) contacting the sample with a probe for HM
expression; (c) comparing the results with a control; and (d) determining
whether the comparison indicates a likelihood disease). (M1) is useful
for identifying modulators of the p53 pathway. A probe for HM expression
is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
in a patient, where the cancer has greater than 25 % expression level.
Modulators identified by (M1) are useful in a variety of diagnostic and
therapeutic applications, where disease or disorder prognosis is related
to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
proliferation disorders (e.g. cancer). Another two new methods (M2 and
M3) are useful for modulating the p53 pathway of a cell, thus restoring
the p53 function of the cell, so that the cell undergoes normal
proliferation or progression through the cell cycle. (M2) and (M3) are
also useful for treating defects in the p53 pathway such as angiogenic,
apoptotic or cell proliferation disorders. The present sequence
represents a human p53 pathway modifying protein
XX
SQ Sequence 660 AA;
Query Match 59.3%; Score 48; DB 6; Length 660;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      2 SLOEFLSHGGYVC 14
Db      454 SLEDFLYHEGYAC 466

RESULT 13
ID      ADL13058
XX      ADL13058 standard; protein; 660 AA.
AC      ADL13058;
XX      06-MAY-2004 (first entry)
XX      Human steroid-induced C3A liver cell protein #121.
XX      Hepatotropic; Gene therapy; Wilson disease; liver disorder;
XX      steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX      Homo sapiens.
XX      US6673549-B1.
XX      06-JAN-2004.
XX      12-OCT-2001; 2001US-00976594.
XX      12-OCT-2000; 2000US-0240409P.
XX      (INCY-) INCYTE CORP.
XX      Furness LM, Buchbinder JL;
XX      WPI; 2004-068610/07.
XX      Combination useful for preparing a composition for treating liver
XX      disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
XX      comprising cDNAs that are differentially expressed in response to steroid
XX      treatment.
XX      Disclosure; SEQ ID NO 787; 141pp; English.
XX      The invention relates to a combination comprising cDNAs that are
XX      differentially expressed in response to steroid treatment. Also included
XX      are the following: a high throughput method for using a cDNA to detect
XX      differential expression of nucleic acids in a sample; and a high
XX      throughput method of screening molecules or compounds to identify a
XX      ligand that specifically binds a cDNA. The sample is from a subject with
XX      Wilson disease and comparison of a standard defines a stage of that
XX      disease. The high throughput method of screening molecules or compounds
XX      to identify a ligand that specifically binds a cDNA comprises: combining
XX      the combination with molecules or compounds under conditions to allow
XX      specific binding; and detecting specific binding between each cDNA and at
XX      least one molecule or compound. The molecules or compounds are regulatory
XX      proteins. The combination is useful for preparing a composition for
XX      treating liver disorders associated with steroid therapy, e.g., cirrhosis
XX      or hepatitis. The present sequence represents a human protein which is
XX      differentially expressed in steroid-induced C3A liver cells. Note: The
XX      sequence data for this patent did not form part of the printed
XX      specification but was obtained in electronic format directly from USPTO
XX      at seqdata.uspto.gov/sequence.html.
XX      Sequence 660 AA;
XX      Query Match      59.3%; Score 48; DB 8; Length 660;
XX      Best Local Similarity 61.5%; Pred. No. 41;
XX      Matches      8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 SLOEFLSHGGYVC 14
Db      454 SLEDFLYHEGYAC 466

RESULT 14
XX      ADL13058 standard; protein; 660 AA.
XX      01-JUL-2004 (first entry)
XX      Human PRO84194 protein SEQ ID NO:10.
XX      human; PRO; antipsoriatic; gene therapy; psoriasis.
XX      Homo sapiens.
XX      WO2004028447-A2.
XX      08-APR-2004.
XX      28-AUG-2003; 2003WO-US027382.
XX      26-SEP-2002; 2002US-0414484P.
XX      (GETH ) GENENTECH INC.
XX      Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX      Wu TD;
XX      WPI; 2004-305092/28.
XX      N-PSDB; ADM33392.
XX      New PRO nucleic acid or polypeptide e.g., PRO37523, PRO71267, PRO71295 or
XX      PRO1843, useful for preparing a composition for diagnosing or treating
XX      psoriasis in a mammal.
XX      Claim 9; SEQ ID NO 10; 143pp; English.
XX      The present invention describes a novel human isolated nucleic acid (1)
XX      designated PRON, where n represents a 4 or 5 digit number. Also
XX      described: (1) a vector comprising (1); (2) a host cell comprising the
XX      vector; (3) a process for producing a PRO polypeptide; (4) an isolated
XX      PRO polypeptide; (5) a chimeric molecule comprising the polypeptide fused
XX      to a heterologous amino acid sequence; (6) an antibody which specifically
XX      binds to a PRO polypeptide; (7) a composition of matter comprising the
XX      polypeptide or its agonist or antagonist or antibody and a carrier; (8)
XX      an article of manufacture comprising a container and its label and the
XX      composition of matter, contained within the container, where label on the
XX      container indicates that the composition of matter can be used for
XX      treating an immune related disorder; (9) a method of treating psoriasis
XX      in a mammal; (10) a method for determining the presence of a PRO
XX      polypeptide in a sample suspected of containing the polypeptide; (11) a
XX      method of identifying a compound that inhibits the expression of a gene
XX      encoding a PRO polypeptide from the present invention; (12) a method of
XX      identifying a compound that mimics the activity of a PRO polypeptide from
XX      the present invention; and (13) a method of diagnosing psoriasis in a
XX      mammal. The PRO sequences have antipsoriatic activity, and can be used in
XX      gene therapy. The PRO nucleic acid or polypeptide sequences from the
XX      present invention can be used for preparing a composition for diagnosing
XX      or treating psoriasis. The present sequence represents a human PRO
XX      protein from the present invention.
XX      Sequence 660 AA;
XX      Query Match      59.3%; Score 48; DB 8; Length 660;
XX      Best Local Similarity 61.5%; Pred. No. 41;
XX      Matches      8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      2 SLOEFLSHGGYVC 14
Db      454 SLEDFLYHEGYAC 466

RESULT 15
XX      ADW08765
XX      ADW08765 standard; protein; 660 AA.
```


PS Disclosure; Fig 3A-B; 54pp; English.

XX This represents the amino acid sequence of the Dead Box X (DBX) gene long

CC transcript. The invention relates to genes occurring on the non-

CC recombining region of the human Y chromosome. The sequences fall into two

CC classes: (1) X-homologous DNA which are expressed in many organs, having

CC functional X homologues and (2) testis-specific DNA sequences. Y

CC chromosomal DNA from males with known conditions such as infertility and

CC reduced sperm count can be assessed using the invention to determine

CC whether the condition is associated with or caused by the occurrence of

CC the gene or gene alteration. Candidate inhibitors of the enzymatic

CC activity of the genes can be assessed using in vitro assays

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 2; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGQYVC 14

Db 456 SLEDFLYHEGYAC 468

RESULT 18

AAW81502

ID AAW81502 standard; protein; 662 AA.

XX

AC AAW81502;

XX

DT 02-FEB-1999 (first entry)

DE

DE Dead Box X (DBX) gene short transcript amino acid sequence.

XX

KW Non-recombining region; human; Y chromosome; X homologue; testis; DBX;

KW infertility; sperm; gene alteration; inhibitor; Dead Box X.

XX

OS Homo sapiens.

XX

PN WO9846747-A2.

XX

PD 22-OCT-1998.

XX

PF 10-APR-1998; 98WO-US007115.

XX

PR 11-APR-1997; 97US-0041877P.

XX

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX

PI Lahn BT, Page DC;

XX

DR WPI; 1998-568729/48.

DR

DR N-PSDB; AAV69632.

XX

PT Novel genes in the non-combining region of Y chromosome - useful to

PT diagnose if male infertility or reduced sperm count has a genetic basis.

XX

PS Disclosure; Fig 3A-B; 54pp; English.

XX

CC This represents the amino acid sequence of the Dead Box X (DBX) gene

CC short transcript. The invention relates to genes occurring on the non-

CC recombining region of the human Y chromosome. The sequences fall into two

CC classes: (1) X-homologous DNA which are expressed in many organs, having

CC functional X homologues and (2) testis-specific DNA sequences. Y

CC chromosomal DNA from males with known conditions such as infertility and

CC reduced sperm count can be assessed using the invention to determine

CC whether the condition is associated with or caused by the occurrence of

CC the gene or gene alteration. Candidate inhibitors of the enzymatic

CC activity of the genes can be assessed using in vitro assays

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 2; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGQYVC 14

Db 456 SLEDFLYHEGYAC 468

RESULT 19

AAW96483

ID AAW96483 standard; protein; 662 AA.

XX

AC AAW96483;

XX

DT 12-SEP-2000 (first entry)

DE

DE Human DBX1, an RNA helicase homologue.

XX

KW Topoisomerase III-alpha; interacting protein; TIII-a IP; RNA helicase;

KW DBX1; nucleosome; supercoiling; chromosome segregation; recombination;

KW stabilization; cell division; apoptosis; cell cycle regulation;

KW cytotoxic; anti-tumour; DBX1.

XX

OS Homo sapiens.

XX

PN WO2000032768-A1.

XX

PD 08-JUN-2000.

XX

PF 29-NOV-1999; 99WO-FR002952.

XX

PR 30-NOV-1998; 98FR-00015081.

XX

PA (AVET) AVENTIS PHARMA SA.

XX

PI Fournier A, Goulaouic H, Riou J;

XX

DR WPI; 2000-412316/35.

DR

DR N-PSDB; AAA29207.

XX

PT New nucleic acid encoding ligand for topoisomerase III alpha, useful for

PT inhibiting the enzyme and in drug screening, e.g. for potential

PT anticancer agents.

XX

PS Disclosure; Page 53-56; 68pp; French.

XX

CC This protein is DBX1, which shows homology with RNA helicases but the

CC activity of a helicase has never been demonstrated and its function has

CC not yet been identified. DBX1 possesses the 8 characteristic motifs of

CC the helicases of the family "DEAD". In particular, it appears to be part

CC of the sub-family represented by helicase PL10. The DBX1 gene is situated

CC on the X chromosome and its homologue, which is situated on the Y

CC chromosome has 91 percent identity with a novel protein coding sequence.

CC The novel protein (see AAW96482) is a human topoisomerase III-alpha

CC interacting protein (TIII-a IP). TIII-a IP has structural features in

CC common with RNA helicases (e.g. DBX1), which are involved in destruction

CC of the nucleosome structure, supercoiling of DNA, segregation of newly

CC replicated chromosomes, and recombination and stabilization of the

CC genome. Agents that alter interaction between TIII-a IP and TIII-a

CC therefore modulate cell division, replication or repair, so may slow cell

CC translation, splicing and DNA recombination or repair, so may slow cell

CC growth, block the cell cycle or induce apoptosis. Antibodies and ligands

CC of TIII-a IP are used to prevent, treat or alleviate diseases that

CC involve abnormal regulation of the cell cycle, i.e. they are potential

CC anti-tumour (cytotoxic) agents

XX

SQ Sequence 662 AA;

Query Match 59.3%; Score 48; DB 3; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGQYVC 14

Db 456 SLEDFLYHEGYAC 468
||:|||||
RESULT 20
ABBS7045
ID ABB57045 standard; protein; 662 AA.
XX
AC ABB57045;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:68.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
WPI; 2002-034733/04.
DR N-PSDB; ABI99243.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 215-218; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (i) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (i). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 662 AA;
Query Match 59.3%; Score 48; DB 5; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLOEFLSHGGYVC 14
||:|||||
Db 456 SLEDFLYHEGYAC 468
||:|||||
RESULT 21
ABO07236
ID ABO07236 standard; protein; 662 AA.
XX
AC ABO07236;
XX
XX

DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 196.
XX
KW Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200299122-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
WPI; 2003-156859/15.
DR N-PSDB; ACD13409.
XX
PT Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
PS Example 2; Page 565-567; 678pp; English.
XX
CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
SQ Sequence 662 AA;
Query Match 59.3%; Score 48; DB 6; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLOEFLSHGGYVC 14
||:|||||

Db 456 SLEDFLYHEGYAC 468

RESULT 22

ABO52958

ID ABO52958 standard; protein; 662 AA.

AC ABO52958;

XX

XX

XX 09-OCT-2003 (first entry)

DT

DE Human spliceosome associated protein (SAP) #64.

DE

XX Human; SAP; spliceosome associated protein; ribonucleoprotein;

KW RNP complex; RNA affinity substrate; RNP assembly sequence;

KW spliceosomal complex; hnRNP complex; mRNA export complex;

KW mRNA localisation complex; RNA editing complex; intron complex;

KW H complex; telomerase complex; fragile X protein complex;

KW reverse transcriptase complex; gene splicing complex.

XX

OS Homo sapiens.

XX

XX US2003068803-A1.

PN

XX

XX 10-APR-2003.

PD

XX

XX 14-JAN-2002; 2002US-00047991.

PF

XX

XX 12-JAN-2001; 2001US-0261521P.

PR

XX (REED//) REED R.

PA (ZHOU//) ZHOU Z.

XX

XX

PI Reed R, Zhou Z;

XX

XX WPI; 2003-540885/51.

DR

XX

XX Isolating ribonucleoprotein complex, by contacting RNA affinity substrate having ribonucleoprotein assembly sequence and affinity tag, with protein mixture, subjecting complex formed to chromatography, affinity selection.

PT

XX

XX Claim 24; Page; 39pp; English.

PS

XX The invention relates to forming (M1) an isolated ribonucleoprotein (RNP) complex (C), involves contacting an RNA affinity substrate (S) comprising an RNP assembly sequence (AS) and an affinity tag, with a protein mixture to permit formation of (C) on AS, subjecting (C) to chromatographic separation, and subjecting (C) to affinity selection, where the affinity tag (e.g. bacteriophage MS2 coat protein in a fusion protein with E. coli maltose binding protein) binds to an affinity matrix. Also included are an isolated spliceosome preparation (isolated by (M1)), a RNA comprising an RNP complex binding site and at least one phase coat protein recognition site, a nucleic acid encoding the RNA, and treating (M2) a subject having a disorder associated with abnormal RNP complexes (by obtaining a sample of cells from a subject, purifying RNP complexes from the cells of the subject by (M1), determining the presence in the purified RNP complexes of one or more proteins, and normalising the amount of RNPs in the subject. (M1) is useful for forming an isolated RNP complex selected from a spliceosomal complex (selected from E, A, B and C complex), an hnRNP complex, an mRNA export complex, an mRNA localisation complex, an RNA editing complex, an intron complex, or an H complex. (M1) is useful in a diagnostic assay for determining whether a subject has abnormal RNP complexes, (M2) is useful for treating a subject having a disorder associated with abnormal RNP complexes. (M1) is useful for forming an isolated RNP complex such as a telomerase complex, a fragile X protein complex, a reverse transcriptase complex or a gene splicing complex. The present sequence represents a known human spliceosome associated protein (SAP) isolated by the methods of the invention. Note: The present sequence is not shown in the specification but was obtained from Genbank or Swissprot using the information provided in table 1 of the specification

XX

XX Sequence 662 AA;

Query Match 59.3%; Score 48; DB 6; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14

Db 456 SLEDFLYHEGYAC 468

||:|:|:|:|:|

RESULT 23

ADN04580

ID ADN04580 standard; protein; 662 AA.

XX

XX ADN04580;

AC

XX 01-JUL-2004 (first entry)

DT

XX Antipsoriatic protein sequence #479.

DE

XX antipsoriatic; gene therapy; psoriasis; diagnosis.

KW

XX Homo sapiens.

XX

XX OS

XX WO2004028479-A2.

PN

XX

XX 08-APR-2004.

PD

XX

XX 25-SEP-2003; 2003WO-US030907.

PF

XX

XX 25-SEP-2002; 2002US-0414006P.

PR

XX (GETH) GENENTECH INC.

PA

XX

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

PI Wu TD;

XX

XX WPI; 2004-305105/28.

DR

XX N-PSDB; ADN04579.

DR

XX New PRO nucleic acid or polypeptide, useful for preparing a pharmaceutical composition for diagnosing or treating psoriasis in a mammal.

PT

XX

XX Claim 9; SEQ ID NO 974; 3069pp; English.

PS

XX The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polypeptides of the invention.

CC

XX Sequence 662 AA;

CC

XX

Query Match 59.3%; Score 48; DB 8; Length 662;

Best Local Similarity 61.5%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14

Db 456 SLEDFLYHEGYAC 468

||:|:|:|:|:|

RESULT 24

ADN03714

ID ADN03714 standard; protein; 662 AA.

XX

XX ADN03714;

AC

XX

XX 01-JUL-2004 (first entry)

DT

XX

XX Antipsoriatic protein sequence #54.

DE

XX

KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
OS WO2004028479-A2.
XX 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
XX (GETH) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN03713.
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 9; SEQ ID NO 108; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX Sequence 662 AA;
SQ
Query Match 59.3%; Score 48; DB 8; Length 662;
Best Local Similarity 61.5%; Pred. NO. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLQEFSLHGGYVC 14
||:|||||
DB 456 SLEDFLYHEGYAC 468
RESULT 25
ADX07534
ID ADX07534 standard; protein; 662 AA.
XX ADX07534;
XX 21-APR-2005 (first entry)
XX Cyclin-dependent kinase modulation biomarker SEQ ID NO 2099.
XX cytotatic; cyclin-dependent kinase; cdk; biomarker.
XX Homo sapiens.
XX WO2005012875-A2.
XX 10-FEB-2005.
XX 29-JUL-2004; 2004WO-US024424.
XX 29-JUL-2003; 2003US-0490890P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
PI WPI; 2005-163068/17.
XX N-PSDB; ADX07533.
XX Biomarkers useful for predicting or determining the response of a mammal
PT
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX Claim 5; SEQ ID NO 2099; 141pp; English.
XX This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO.1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
CC oxazolyl)methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published pct sequences. This
CC sequence represents a biomarker used in the method of the invention.
XX Sequence 662 AA;
SQ
Query Match 59.3%; Score 48; DB 9; Length 662;
Best Local Similarity 61.5%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 SLQEFSLHGGYVC 14
||:|||||
DB 456 SLEDFLYHEGYAC 468
RESULT 26
AAG07259
ID AAG07259 standard; protein; 192 AA.
XX AAG07259;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 4345.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135622P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138884P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0144815P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158039P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 18-OCT-1999; 99US-0159585P.

CC	plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC	exemplification of the invention.
XX	
SQ	Sequence 198 AA;
	Query Match 55.6%; Score 45; DB 8; Length 198;
	Best Local Similarity 72.7%; Pred. No. 37;
	Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY	3 LQEFLSHGYY 13
	:
Db	116 LQEFVNHGGYY 126
RESULT 28	
AAG05910	
ID	AAG05910 standard; protein; 251 AA.
XX	AC AAG05910;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 2491.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	28-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138884P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145114P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145591P.
PR 27-JUL-1999; 99US-0145591P.
PR 27-JUL-1999; 99US-0145591P.
PR 28-JUL-1999; 99US-0145591P.
PR 28-JUL-1999; 99US-0145591P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149388P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 45; DB 3; Length 251;
Best Local Similarity 50.0%; Pred. No. 47;

Matches	7;	Conservative	4;	Mismatches	3;	Indels	0;	Gaps	0;
QY	1	CSLQEFLSHGGVVC 14							
DB	28	CALKWFFVNHGTVVC 41							
RESULT 29									
AAYS9428									
ID	AAYS9428 standard; protein; 313 AA.								
XX									
AC	AAYS9428;								
XX									
DT	21-MAR-2000 (first entry)								
XX									
DE	Soybean inositol 1,3,4-triphosphate 5/6-kinase protein sequence.								
XX									
KW	Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;								
KW	myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;								
KW	animal feed.								
XX									
OS	Glycine max.								
XX									
PN	WO9955879-A1.								
XX									
PD	04-NOV-1999.								
XX									
PF	22-APR-1999; 99WO-US008790.								
XX									
PR	24-APR-1998; 98US-0082960P.								
XX									
PA	(DUPO) DU PONT DE NEMOURS & CO E I.								
XX									
PI	Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;								
XX									
DR	WPI: 2000-072179/06.								
DR	N-PSDB; AA248812.								
XX									
PT	Novel phytic acid biosynthetic enzymes used to alter the level of the								
PT	enzyme in transgenic plants.								
XX									
PS	Claim 5; Page 53-54; 63pp; English.								
XX									
CC	This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein								
CC	of the invention. ITK is a phytic acid (Myo-inositol 1,2,3,4,5,6-								
CC	hexaphosphate) biosynthetic enzyme. The ITK enzymes of the invention may								
CC	be prepared recombinantly and used to raise antibodies, which are used								
CC	for detecting the enzymes in situ in cells or in vitro in cell extracts.								
CC	The polynucleotides of the invention may be used to create transgenic								
CC	plants in which the ITK levels are present at higher or lower levels than								
CC	normal, or in cell types or developmental processes where they are not								
CC	normally found. This would alter the level of 1,3,4-triphosphate 5/6								
CC	kinase found in those cells. In addition, it may be desirable to								
CC	eliminate expression of ITK genes for certain applications. The								
CC	polynucleotides also serve as a source of probes and primers, which are								
CC	useful for genetic mapping, as markers for traits linked to those genes,								
CC	and to isolate homologous sequences from other species, as well as for								
CC	physical mapping. Decreased amounts of phytate and increased amounts of								
CC	available phosphate in animal feeds such as corn would lead to improved								
CC	feed efficiency. The proteins of the present invention lead to a better								
CC	understanding of the phytic acid biosynthesis pathway, allowing it to be								
CC	exploited for commercial uses, e.g. in animal feeds								
XX									
SQ	Sequence 313 AA;								
Query Match 55.6%; Score 45; DB 3; Length 313;									
Best Local Similarity 72.7%; Pred. No. 59;									
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;									
QY	3	LQEFSLHGGVY 13							
DB	177	LQEFVNHGGV 187							

RESULT 30	
AAG05909	
ID	AAG05909 standard; protein; 370 AA.
XX	
AC	AAG05909;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 2490.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	16-APR-1999; 99US-0128714P.
PR	19-APR-1999; 99US-0129845P.
PR	21-APR-1999; 99US-0130077P.
PR	23-APR-1999; 99US-0130443P.
PR	28-APR-1999; 99US-0130510P.
PR	30-APR-1999; 99US-0130891P.
PR	04-MAY-1999; 99US-0131449P.
PR	05-MAY-1999; 99US-0132048P.
PR	06-MAY-1999; 99US-0132407P.
PR	06-MAY-1999; 99US-0132484P.
PR	07-MAY-1999; 99US-0132486P.
PR	11-MAY-1999; 99US-0132863P.
PR	14-MAY-1999; 99US-0134258P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	18-MAY-1999; 99US-0134370P.
PR	19-MAY-1999; 99US-0134768P.
PR	20-MAY-1999; 99US-0134941P.
PR	21-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 55.6%; Score 45; DB 3; Length 370;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLOEFLSHGGYVC 14
|::|::|::|::|
Db 147 CALKWFVNHGSTVC 160

RESULT 31
ABU27233
1D ABU27233 standard; protein; 536 AA.
XX
AC ABU27233;
XX

DT 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #12760.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Chlamydia trachomatis.
KW WO200277183-A2.
OS 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI WPI; 2003-029926/02.
DR N-PSDB; ACA311103.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 55157; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC or a gene on which the test compound that influences the activity of
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 536 AA;
SQ Query Match 55.6%; Score 45; DB 6; Length 536;
Best Local Similarity 63.6%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 2 SLOEFLSHGGY 12

Db 512 SIKEFLAAGGH 522
RESULT 32
AAU23148
ID AAU23148 standard; protein; 179 AA.
XX AC AAU23148;
XX DT 17-DEC-2001 (first entry)
XX DE Novel human enzyme polypeptide #234.
XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX OS Homo sapiens.
XX PN WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001239.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.

```
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
XX N-PSDB; AAS41018.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
XX treating neural, immune system, muscular, reproductive, pulmonary,
XX cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 11; SEQ ID NO 1144; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
XX polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
XX encoding them. The enzyme polypeptides of the invention may comprise the
XX functional classes of oxidoreductases, transferases, hydrolases, lyases,
XX isomerases or ligases. The sequences of the invention are useful in the
XX diagnosis, treatment, prevention and/or prognosis of a wide range of
XX disorders, including hyperproliferative disorders (e.g. cancer),
XX immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
XX arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
XX disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
XX cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
XX (e.g. haemophilia), reproductive disorders (e.g. infertility) and
XX infectious disorders (e.g. Influenza). The polynucleotides of the
XX invention can also be used in gene therapy. AAU22915-AAU23814 represent
XX the novel human enzyme polypeptides of the invention. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pot_sequences
XX
XX Sequence 179 AA;
XX
XX Query Match 54.3%; Score 44; DB 4; Length 179;
XX Best Local Similarity 50.0%; Pred. No. 49;
XX Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
XX
XX Qy 1 CSLQFLSHGGYVC 14
XX | | | | | : | |
XX Db 62 CPLPPMVSHGDFVC 75
XX
XX RESULT 33
XX AAU87567
XX ID AAU87567 standard; protein; 179 AA.
XX
XX AC AAU87567;
XX
XX AC AAU87567;
XX
XX 05-JUN-2002 (first entry)
XX
XX DE Novel central nervous system protein #477.
XX
```


Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
hyperproliferative disorder; neoplasm; cardiovascular disorder;
cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
adenocarcinoma; reproductive system disorder; testicular feminisation;
endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
respiratory disorder; renal disorder; kidney failure; blood disorder;
myocardial infarction; wound healing; cell proliferation; skin aging;
food additive; food preservative; gene therapy.

KW	PR	14-SEP-2000;	2000US-0232399P.
KW	PR	14-SEP-2000;	2000US-0232400P.
KW	PR	14-SEP-2000;	2000US-0232401P.
KW	PR	14-SEP-2000;	2000US-0233063P.
KW	PR	14-SEP-2000;	2000US-0233064P.
KW	PR	14-SEP-2000;	2000US-0233065P.
KW	PR	21-SEP-2000;	2000US-0234223P.
KW	PR	21-SEP-2000;	2000US-0234274P.
KW	PR	25-SEP-2000;	2000US-0234997P.
KW	PR	25-SEP-2000;	2000US-0234998P.
KW	PR	26-SEP-2000;	2000US-0235484P.
XX	PR	27-SEP-2000;	2000US-0235834P.
OS	PR	27-SEP-2000;	2000US-0235836P.
XX	PR	29-SEP-2000;	2000US-0236327P.
PN	PR	29-SEP-2000;	2000US-0236367P.
XX	PR	29-SEP-2000;	2000US-0236368P.
PD	PR	29-SEP-2000;	2000US-0236370P.
XX	PR	02-OCT-2000;	2000US-0236802P.
XX	PR	02-OCT-2000;	2000US-0237037P.
PR	PR	02-OCT-2000;	2000US-0237038P.
PR	PR	02-OCT-2000;	2000US-0237039P.
PR	PR	02-OCT-2000;	2000US-0237040P.
PR	PR	13-OCT-2000;	2000US-0239935P.
PR	PR	13-OCT-2000;	2000US-0239937P.
PR	PR	20-OCT-2000;	2000US-0240960P.
PR	PR	20-OCT-2000;	2000US-0241221P.
PR	PR	20-OCT-2000;	2000US-0241785P.
PR	PR	20-OCT-2000;	2000US-0241786P.
PR	PR	20-OCT-2000;	2000US-0241787P.
PR	PR	20-OCT-2000;	2000US-0241808P.
PR	PR	20-OCT-2000;	2000US-0241809P.
PR	PR	20-OCT-2000;	2000US-0241826P.
PR	PR	01-NOV-2000;	2000US-0244617P.
PR	PR	08-NOV-2000;	2000US-0246474P.
PR	PR	08-NOV-2000;	2000US-0246475P.
PR	PR	08-NOV-2000;	2000US-0246476P.
PR	PR	08-NOV-2000;	2000US-0246477P.
PR	PR	08-NOV-2000;	2000US-0246478P.
PR	PR	08-NOV-2000;	2000US-0246523P.
PR	PR	08-NOV-2000;	2000US-0246524P.
PR	PR	08-NOV-2000;	2000US-0246525P.
PR	PR	08-NOV-2000;	2000US-0246526P.
PR	PR	08-NOV-2000;	2000US-0246527P.
PR	PR	08-NOV-2000;	2000US-0246528P.
PR	PR	08-NOV-2000;	2000US-0246532P.
PR	PR	08-NOV-2000;	2000US-0246609P.
PR	PR	08-NOV-2000;	2000US-0246610P.
PR	PR	08-NOV-2000;	2000US-0246611P.
PR	PR	08-NOV-2000;	2000US-0246613P.
PR	PR	17-NOV-2000;	2000US-0249207P.
PR	PR	17-NOV-2000;	2000US-0249208P.
PR	PR	17-NOV-2000;	2000US-0249209P.
PR	PR	17-NOV-2000;	2000US-0249210P.
PR	PR	17-NOV-2000;	2000US-0249211P.
PR	PR	17-NOV-2000;	2000US-0249212P.
PR	PR	17-NOV-2000;	2000US-0249213P.
PR	PR	17-NOV-2000;	2000US-0249214P.
PR	PR	17-NOV-2000;	2000US-0249215P.
PR	PR	17-NOV-2000;	2000US-0249216P.
PR	PR	17-NOV-2000;	2000US-0249217P.
PR	PR	17-NOV-2000;	2000US-0249218P.
PR	PR	17-NOV-2000;	2000US-0249244P.
PR	PR	17-NOV-2000;	2000US-0249245P.
PR	PR	17-NOV-2000;	2000US-0249264P.
PR	PR	17-NOV-2000;	2000US-0249265P.
PR	PR	17-NOV-2000;	2000US-0249297P.
PR	PR	17-NOV-2000;	2000US-0249299P.
PR	PR	17-NOV-2000;	2000US-0249300P.
PR	PR	01-DEC-2000;	2000US-0250160P.
PR	PR	01-DEC-2000;	2000US-0250391P.
PR	PR	05-DEC-2000;	2000US-0251030P.
PR	PR	05-DEC-2000;	2000US-0251988P.

05-DEC-2000; 2000US-0256719P.
 06-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-581633/65.
 N-PSDB; ABK43897.
 New isolated nucleic acid encoding a protein for diagnosing, preventing,
 treating or ameliorating medical conditions and used as food additives or
 preservatives.
 Claim 9; SEQ ID NO 1085; 837pp; English.
 The invention describes an isolated nucleic acid molecule (I) encoding a
 novel central nervous system protein. (I) and polypeptides (III) encoded
 by (I), are used to treat a medical conditions and in diagnosis of a
 pathological condition. Disorders which are diagnosed or treated include
 autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 angiodysplasia, nervous system disorders e.g. Alzheimer's disease and
 amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 adenocarcinomas and irritable bowel syndrome, reproductive system
 disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 leukaemia, disorders involving neovascularisation e.g. malignancies,
 respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 acute kidney failure and blood related disorders e.g. myocardial
 infarction. The polypeptides can also be used to aid wound healing and
 epithelial cell proliferation, to prevent skin aging due to sunburn, to
 maintain organs before transplantation, for supporting cell culture of
 primary tissues, to regenerate tissues and in chemotaxis. The
 polypeptides can also be used as a food additive or preservative to
 increase or decrease storage capabilities, fat content, lipid, protein,
 Query Match 54.3%; Score 44; DB 4; Length 179;
 Best Local Similarity 50.0%; Pred. No. 49;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 CSLOEPLSHGQYVC 14
 Db 62 CPLEPMVSHGDFVC 75
 RESULT 34
 ADI54882
 ID ADI54882 standard; protein; 179 AA.
 AC ADI54882;
 XX ADI54882;
 DT 22-APR-2004 (first entry)
 XX Novel human protein seq id 1085.
 XX neuroprotective; neurotropic; antiparkinsonian; anticonvulsant;
 KW antidiabetic; antirheumatic; antiarthritic; dermatologic;
 KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
 KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
 KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
 KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;
 KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;

KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
 KW amyotrophic lateral sclerosis; multiple sclerosis;
 KW immune system disorder; diabetes; rheumatoid arthritis;
 KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
 KW inflammatory bowel disease; ischaemia-reperfusion injury;
 KW HIV infection; hepatitis infection; Crohn's disease; infectious disease;
 KW fungal infection; parasitic infection; bacterial infection;
 KW reproductive disorder; gastrointestinal disorder; muscular disorder;
 KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
 KW renal disorder; acute glomerulonephritis; pyelonephritis;
 KW renal lithiasis; proliferative disorder; cancerous diseases; human.
 XX Homo sapiens.
 OS
 XX US2004018969-A1.
 XX
 XX 29-JAN-2004.
 XX
 XX 17-JAN-2001; 2001US-00764875.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.

Mon May 15 11:35:36 2006

```
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-02324063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234221P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2004-122079/12.
DR N-PSDB; ADI54284.
XX
XX New polypeptides and nucleic acid molecules, useful for detecting,
PT preventing, diagnosing, prognosticating, treating or ameliorating medical
PT conditions e.g. neural disorders, reproductive disorders or infectious
PT diseases.
XX
PS Claim 11; SEQ ID NO 1085; 413pp; English.
XX
XX The invention describes an isolated polypeptide comprising an amino acid
CC sequence at least 90% identical to: a polypeptide fragment, domain,
CC epitope, or full-length protein of any one of 607 amino acid sequences
CC (I) described in the specification; a polypeptide fragment of (I), or the
CC encoded sequence contained in (II), having biological activity; or a
CC variant, allelic variant, or a species homologue of (I). The polypeptides
CC and nucleic acid molecules are useful for detecting, preventing,
CC diagnosing, prognosticating, treating or ameliorating medical conditions
CC such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
CC Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
CC immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic

Query Match 54.3%; Score 44; DB 8; Length 179;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
DB 62 CPLPPMVSHGDFVC 75

RESULT 35
ABG26613
ID ABG26613 standard; protein; 184 AA.
XX
AC ABG26613;
XX
XX 18-FEB-2002 (first entry)
DT
XX
DE Novel human diagnostic protein #26604.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
XX WO2001/5067-A2.
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
```

```
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS90800.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 56972; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: the sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 184 AA;
XX
XX Query Match 54.3%; Score 44; DB 4; Length 184;
XX Best Local Similarity 50.0%; Pred. NO. 50;
XX Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
XX
XX Qy 1 CSLQEFLSHGQYVC 14
XX | | : | | : | |
XX Db 113 CPLPPMVSHGDFVC 126
XX
XX RESULT 36
XX ABB03078
XX ID ABB03078 standard; protein; 187 AA.
XX AC ABB03078;
XX XX
XX DT 07-JAN-2002 (first entry)
XX XX
XX DE Human expressed polypeptide SEQ ID NO 51.
XX
XX Cystostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
XX
XX WO200155167-A1.
XX PN
XX XX
XX PD 02-AUG-2001.
XX PF
XX 17-JAN-2001; 2001WO-US001319.
XX XX
XX 31-JAN-2000; 2000US-0175065P.
XX PR
XX 04-FEB-2000; 2000US-0180628P.
XX
```

```
PR 26-SEP-2000; 2000US-0235484P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465559/50.
XX N-PSDB; AAI99630.
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
XX diagnose diseases or disorders associated with polypeptides and for
XX treating autoimmune diseases e.g., multiple sclerosis, rheumatoid
XX arthritis.
XX
XX Claim 11; SEQ ID NO 51; 504pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAI99614-AAI99654) and proteins
XX (ABB03062-ABB03085) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 187 AA;
XX
XX Query Match 54.3%; Score 44; DB 4; Length 187;
XX Best Local Similarity 50.0%; Pred. No. 51;
XX Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
XX
XX Qy 1 CSLQEFLSHGQYVC 14
XX | | : | | : | |
XX Db 62 CPLPPMVSHGDFVC 75
XX
XX RESULT 37
XX AAU87270
XX ID AAU87270 standard; protein; 187 AA.
XX XX
XX AC AAU87270;
XX XX
XX DT 05-JUN-2002 (first entry)
XX XX
XX DE Novel central nervous system protein #180.
XX
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
XX hyperproliferative disorder; neoplasm; cardiovascular disorder;
XX cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
XX nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
XX acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
XX adenocarcinoma; reproductive system disorder; testicular feminisation;
XX endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
XX respiratory disorder; renal disorder; kidney failure; blood disorder;
XX myocardial infarction; wound healing; cell proliferation; skin aging;
XX food additive; food preservative; gene therapy.
XX
XX Homo sapiens.
XX
```

PN WO200155318-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001332.
XX 31-JAN-2000; 2000US-0179065P.
PD 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217496P.
XX 11-JUL-2000; 2000US-0218290P.
XX 14-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226688P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0228924P.
XX 30-AUG-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 26-SEP-2000; 2000US-0234984P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX PI

XX WPI; 2001-581633/65.
DR N-PSDB; ABK43600.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
XX Claim 9; SEQ ID NO 788; 837pp; English.
PS
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Query Match 54.3%; Score 44; DB 4; Length 187;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 CSLQEFLSHGQYVC 14
| | | | | : | |
| | | | | : | |
Db 62 CPLPMVSHGDFVC 75
RESULT 38
AD154585
ID AD154585 standard; protein; 187 AA.
XX
XX AC AD154585;
XX
DT 16-DEC-2004 (first entry)
XX
XX DE Novel human protein seq id 788.
XX
XX neuroprotective; neurotropic; antiparkinsonian; anticonvulsant;
KW antidiabetic; antirheumatic; antiarthritic; dermatological;
KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
KW cardiovascular; antiarteriosclerotic; antiarrhythmic; cardiac;
KW nephrotropic; litholytic; cytostatic; gene therapy; neural disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
KW amyotrophic lateral sclerosis; multiple sclerosis;
KW immune system disorder; diabetes; rheumatoid arthritis;
KW systemic lupus erythematosus; autoimmune thyroiditis; haemolytic anaemia;
KW inflammatory disorder; ischaemia-reperfusion injury;
KW inflammatory bowel disease; Crohn's disease; infectious disease;
KW HIV infection; hepatitis infection; bacterial infection;
KW fungal infection; parasitic infection; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
KW renal disorder; acute glomerulonephritis; pyelonephritis;
KW renal lithiasis; proliferative disorder; cancerous diseases; human.

OS Homo sapiens.
XX
XX US2004018969-A1.
PD
XX 29-JAN-2004.
XX
PF 17-JAN-2001; 2001US-00764875.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.

PR	27-SEP-2000;	2000US-0235834P.	PA	(RUBE/) RUBEN S M.
PR	27-SEP-2000;	2000US-0235836P.	PA	(BARA/) BARASH S C.
PR	29-SEP-2000;	2000US-0236327P.	XX	
PR	29-SEP-2000;	2000US-0236367P.	PI	Rosen CA, Ruben SM, Barash SC;
PR	29-SEP-2000;	2000US-0236368P.	XX	
PR	29-SEP-2000;	2000US-0236369P.	DR	WPI; 2004-122079/12.
PR	29-SEP-2000;	2000US-0236370P.	DR	N-PSDB; ADI53987.
PR	29-SEP-2000;	2000US-0236802P.	XX	
PR	02-OCT-2000;	2000US-0237037P.	PT	New polypeptides and nucleic acid molecules, useful for detecting,
PR	02-OCT-2000;	2000US-0237038P.	PT	preventing, diagnosing, prognosticating, treating or ameliorating medical
PR	02-OCT-2000;	2000US-0237039P.	PT	conditions e.g. neural disorders, reproductive disorders or infectious
PR	02-OCT-2000;	2000US-0237040P.	PT	diseases.
PR	13-OCT-2000;	2000US-0239935P.	XX	
PR	13-OCT-2000;	2000US-0239937P.	XX	Claim 11; SEQ ID NO 788; 413pp; English.
PR	20-OCT-2000;	2000US-0240960P.	PS	
PR	20-OCT-2000;	2000US-0241221P.	XX	
PR	20-OCT-2000;	2000US-0241785P.	CC	The invention describes an isolated polypeptide comprising an amino acid
PR	20-OCT-2000;	2000US-0241786P.	CC	sequence at least 90% identical to: a polypeptide fragment, domain,
PR	20-OCT-2000;	2000US-0241787P.	CC	epitope, or full-length protein of any one of 607 amino acid sequences
PR	20-OCT-2000;	2000US-0241788P.	CC	(i) described in the specification; a polypeptide fragment of (i), or the
PR	20-OCT-2000;	2000US-0241808P.	CC	encoded sequence contained in (ii), having biological activity; or a
PR	20-OCT-2000;	2000US-0241809P.	CC	variant, allelic variant, or a species homologue of (i). The polypeptides
PR	20-OCT-2000;	2000US-0241810P.	CC	and nucleic acid molecules are useful for detecting, preventing,
PR	01-NOV-2000;	2000US-0244617P.	CC	diagnosing, prognosticating, treating or ameliorating medical conditions
PR	08-NOV-2000;	2000US-0246474P.	CC	such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,
PR	08-NOV-2000;	2000US-0246475P.	CC	Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,
PR	08-NOV-2000;	2000US-0246476P.	CC	immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic
PR	08-NOV-2000;	2000US-0246477P.	CC	
PR	08-NOV-2000;	2000US-0246478P.	CC	
PR	08-NOV-2000;	2000US-0246479P.	CC	
PR	08-NOV-2000;	2000US-0246523P.	CC	
PR	08-NOV-2000;	2000US-0246524P.	CC	
PR	08-NOV-2000;	2000US-0246525P.	CC	
PR	08-NOV-2000;	2000US-0246526P.	CC	
PR	08-NOV-2000;	2000US-0246527P.	CC	
PR	08-NOV-2000;	2000US-0246528P.	CC	
PR	08-NOV-2000;	2000US-0246532P.	CC	
PR	08-NOV-2000;	2000US-0246609P.	CC	
PR	08-NOV-2000;	2000US-0246610P.	CC	
PR	08-NOV-2000;	2000US-0246611P.	CC	
PR	08-NOV-2000;	2000US-0246613P.	CC	
PR	17-NOV-2000;	2000US-0249207P.	PR	
PR	17-NOV-2000;	2000US-0249208P.	PR	
PR	17-NOV-2000;	2000US-0249209P.	PR	
PR	17-NOV-2000;	2000US-0249210P.	PR	
PR	17-NOV-2000;	2000US-0249211P.	PR	
PR	17-NOV-2000;	2000US-0249212P.	PR	
PR	17-NOV-2000;	2000US-0249213P.	PR	
PR	17-NOV-2000;	2000US-0249214P.	PR	
PR	17-NOV-2000;	2000US-0249215P.	PR	
PR	17-NOV-2000;	2000US-0249216P.	PR	
PR	17-NOV-2000;	2000US-0249217P.	PR	
PR	17-NOV-2000;	2000US-0249218P.	PR	
PR	17-NOV-2000;	2000US-0249244P.	PR	
PR	17-NOV-2000;	2000US-0249245P.	PR	
PR	17-NOV-2000;	2000US-0249264P.	PR	
PR	17-NOV-2000;	2000US-0249265P.	PR	
PR	17-NOV-2000;	2000US-0249297P.	PR	
PR	17-NOV-2000;	2000US-0249299P.	PR	
PR	17-NOV-2000;	2000US-0249300P.	PR	
PR	01-DEC-2000;	2000US-0250160P.	PR	
PR	01-DEC-2000;	2000US-0250391P.	PR	
PR	05-DEC-2000;	2000US-0251030P.	PR	
PR	05-DEC-2000;	2000US-0251988P.	PR	
PR	05-DEC-2000;	2000US-0256719P.	PR	
PR	06-DEC-2000;	2000US-0251479P.	PR	
PR	08-DEC-2000;	2000US-0251856P.	PR	
PR	08-DEC-2000;	2000US-0251868P.	PR	
PR	08-DEC-2000;	2000US-0251989P.	PR	
PR	08-DEC-2000;	2000US-0251990P.	PR	
PR	11-DEC-2000;	2000US-0254097P.	PR	
PR	05-JAN-2001;	2001US-0259678P.	PR	
XX			PR	(ROSE/) ROSEN C A.
PA			PR	
<p>Query Match 54.3%; Score 44; DB 8; Length 187;</p> <p>Best Local Similarity 50.0%; Pred. No. 51;</p> <p>Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;</p>				
QY	1	CSLQEFILSHGGYVC 14		
DB	62	CPLPPMVSHGDFVC 75		
<p>RESULT 39</p> <p>AAG44384</p> <p>ID AAG44384 standard; protein; 196 AA.</p> <p>XX AAG44384;</p> <p>AC AAG44384;</p> <p>XX 18-OCT-2000 (first entry)</p> <p>XX Arabidopsis thaliana protein fragment SEQ ID NO: 55591.</p> <p>DE Arabidopsis thaliana.</p> <p>XX Protein identification; signal transduction pathway; metabolic pathway;</p> <p>KW hybridisation assay; genetic mapping; gene expression control; promoter;</p> <p>KW termination sequence.</p> <p>XX Arabidopsis thaliana.</p> <p>OS Arabidopsis thaliana.</p> <p>XX EP1033405-A2.</p> <p>XX 06-SEP-2000.</p> <p>XX 25-FEB-2000; 2000EP-00301439.</p> <p>XX 25-FEB-1999; 99US-0121825P.</p> <p>PR 05-MAR-1999; 99US-0123180P.</p> <p>PR 09-MAR-1999; 99US-0123548P.</p> <p>PR 23-MAR-1999; 99US-0125788P.</p> <p>PR 25-MAR-1999; 99US-0126264P.</p> <p>PR 29-MAR-1999; 99US-0126785P.</p> <p>PR 01-APR-1999; 99US-0127462P.</p> <p>PR 06-APR-1999; 99US-0128234P.</p> <p>PR 08-APR-1999; 99US-0128714P.</p> <p>PR 16-APR-1999; 99US-0129845P.</p> <p>PR 19-APR-1999; 99US-0130077P.</p> <p>PR 21-APR-1999; 99US-0130449P.</p> <p>PR 23-APR-1999; 99US-0130510P.</p> <p>PR 23-APR-1999; 99US-0130891P.</p> <p>PR 28-APR-1999; 99US-0131449P.</p>				

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134286P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135144P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140335P.
PR 23-JUN-1999; 99US-0140336P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0144815P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147600P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.	PR	11-MAY-1999;	99US-0134256P.
PR	21-OCT-1999;	99US-0160767P.	PR	14-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	22-OCT-1999;	99US-0161404P.	PR	21-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161405P.	PR	24-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161406P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161359P.	PR	27-MAY-1999;	99US-0136352P.
PR	26-OCT-1999;	99US-0161360P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161361P.	PR	01-JUN-1999;	99US-0137222P.
PR	28-OCT-1999;	99US-0161920P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161920P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161993P.	PR	07-JUN-1999;	99US-0137724P.
PR	29-OCT-1999;	99US-0162142P.	PR	08-JUN-1999;	99US-0138094P.
Query Match 54.3%; Score 44; DB 3; Length 196;					
Best Local Similarity 63.6%; Pred. No. 53;					
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;					
Qy	3 LOEFLSHGGYV 13				
Db	:				
	53 LOEFVNHGGVI 63				
RESULT 40					
AAG44383					
ID	AAG44383 standard; protein; 222 AA.				
XX					
AC	AAG44383;				
XX					
DT	18-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 55590.				
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-00301439.				
XX					
XX	25-FEB-1999;	99US-0121825P.	PR	13-JUL-1999;	99US-0143542P.
PR	05-MAR-1999;	99US-0123180P.	PR	14-JUL-1999;	99US-0143624P.
PR	09-MAR-1999;	99US-0123548P.	PR	15-JUL-1999;	99US-0144005P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144085P.
PR	25-MAR-1999;	99US-0126264P.	PR	16-JUL-1999;	99US-0144086P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144325P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	99US-0129845P.	PR	20-JUL-1999;	99US-0144352P.
PR	19-APR-1999;	99US-0130077P.	PR	20-JUL-1999;	99US-0144632P.
PR	21-APR-1999;	99US-0130449P.	PR	21-JUL-1999;	99US-0144814P.
PR	23-APR-1999;	99US-0130510P.	PR	21-JUL-1999;	99US-0145086P.
PR	28-APR-1999;	99US-0130891P.	PR	21-JUL-1999;	99US-0145088P.
PR	30-APR-1999;	99US-0131449P.	PR	22-JUL-1999;	99US-0145087P.
PR	30-APR-1999;	99US-0132048P.	PR	22-JUL-1999;	99US-0145089P.
PR	30-APR-1999;	99US-0132407P.	PR	22-JUL-1999;	99US-0145192P.
PR	04-MAY-1999;	99US-0132484P.	PR	23-JUL-1999;	99US-0145145P.
PR	05-MAY-1999;	99US-0132485P.			
PR	06-MAY-1999;	99US-0132486P.			
PR	06-MAY-1999;	99US-0132487P.			
PR	07-MAY-1999;	99US-0132863P.			

```
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145226P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145291P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147361P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149360P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158223P.
PR 12-OCT-1999; 99US-0158363P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 222;
Best Local Similarity 63.6%; Pred. No. 61;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LQEFSLHGGYV 13
Db 79 LQEPVNHGGVI 89

RESULT 41
ADJ32010
ID ADJ32010 standard; protein; 238 AA.
XX AC ADJ32010;
XX DT 15-APR-2004 (first entry)
XX DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #2.
XX KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
XX KW physical mapping; plant breeding; plant; enzyme.
XX OS Unidentified.
XX PN US2003229917-A1.
XX PD 11-DEC-2003.
XX PF 20-DEC-2002; 2002US-00324316.
XX PR 11-OCT-2000; 2000US-00686521.
XX PA (CAHO/) CAHOON R E.
XX PA (HITZ/) HITZ W D.
XX PA (PEAR/) PEARLSTEIN R W.
XX PA (CARL/) CARLSON T J.
XX PI Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WP; 2004-052054/05.
XX DR New polynucleotide comprising a sequence encoding a polypeptide having
PT inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
PT physical mapping for plant breeding.
XX Example 3; Fig 1; 32pp; English.
XX This invention relates to a novel isolated polynucleotide which comprises
CC a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
CC kinase activity and comprising a sequence having 80% identity based on
CC the Clustal alignment method with a fully defined sequence comprising 139
CC amino acids, or its complement. The polynucleotide is useful in genetic
CC or physical mapping for plant breeding. The present sequence is that of a
CC plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
CC exemplification of the invention.
XX Sequence 238 AA;

Query Match 54.3%; Score 44; DB 8; Length 238;
Best Local Similarity 63.6%; Pred. No. 65;
```

```
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LOEFLSHGGYV 13
DB 133 LOEFVNHGGVI 143

RESULT 42
ADJ32009
ID ADJ32009 standard; protein; 238 AA.
XX
AC
XX
XX
DT 15-APR-2004 (first entry)
XX
DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #1.
XX
KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
XX physical mapping; plant breeding; plant; enzyme.
XX
OS Unidentified.
XX
XX US2003229917-A1.
XX
XX 11-DEC-2003.
XX
XX 20-DEC-2002; 2002US-00324316.
XX
XX 11-OCT-2000; 2000US-00686521.
XX
XX (CAHO/) CAHOON R E.
XX (HITZ/) HITZ W D.
XX (PEAR/) PEARLSTEIN R W.
XX (CARL/) CARLSON T J.
XX
XX Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
XX inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
XX physical mapping for plant breeding.
XX
XX Example 3; Fig 1; 32pp; English.
XX
XX This invention relates to a novel isolated polynucleotide which comprises
XX a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
XX kinase activity and comprising a sequence having 80% identity based on
XX the Clustal alignment method with a fully defined sequence comprising 139
XX amino acids, or its complement. The polynucleotide is useful in genetic
XX or physical mapping for plant breeding. The present sequence is that of a
XX plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
XX exemplification of the invention. Note: The SeqID numbers given in figure
XX 1 do not correspond to those in the sequence listing for SeqIDs 2-16.
XX
XX Sequence 238 AA;
Query Match 54.3%; Score 44; DB 8; Length 238;
Best Local Similarity 63.6%; Pred. NO. 65;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LOEFLSHGGYV 13
DB 133 LOEFVNHGGVI 143

RESULT 43
ADJ32014
ID ADJ32014 standard; protein; 247 AA.
XX
AC
XX
XX ADJ32014;
XX
DT 15-APR-2004 (first entry)
```

```
XX DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #6.
XX
XX inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
XX KW physical mapping; plant breeding; plant; enzyme.
XX
XX OS Zea mays.
XX
XX US2003229917-A1.
XX
XX 11-DEC-2003.
XX
XX 20-DEC-2002; 2002US-00324316.
XX
XX 11-OCT-2000; 2000US-00686521.
XX
XX (CAHO/) CAHOON R E.
XX (HITZ/) HITZ W D.
XX (PEAR/) PEARLSTEIN R W.
XX (CARL/) CARLSON T J.
XX
XX Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
XX inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
XX physical mapping for plant breeding.
XX
XX Example 3; Fig 1; 32pp; English.
XX
XX This invention relates to a novel isolated polynucleotide which comprises
XX a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
XX kinase activity and comprising a sequence having 80% identity based on
XX the Clustal alignment method with a fully defined sequence comprising 139
XX amino acids, or its complement. The polynucleotide is useful in genetic
XX or physical mapping for plant breeding. The present sequence is that of a
XX plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
XX exemplification of the invention.
XX
XX Sequence 247 AA;
Query Match 54.3%; Score 44; DB 8; Length 247;
Best Local Similarity 63.6%; Pred. NO. 68;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 LOEFLSHGGYV 13
DB 143 LOEFVNHGGVI 153

RESULT 44
ADJ32016
ID ADJ32016 standard; protein; 261 AA.
XX
XX ADJ32016;
XX
XX 15-APR-2004 (first entry)
XX
XX Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #8.
XX
XX inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
XX KW physical mapping; plant breeding; plant; enzyme.
XX
XX OS Oryza sativa.
XX
XX US2003229917-A1.
XX
XX 11-DEC-2003.
XX
XX 20-DEC-2002; 2002US-00324316.
XX
XX 11-OCT-2000; 2000US-00686521.
XX
```

```
XX (CAHO/) CAHOON R E.
PA (HITZ/) HITZ W D.
PA (PEAR/) PEARLSTEIN R W.
PA (CARL/) CARLSON T J.
XX
XX Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
XX inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
XX physical mapping for plant breeding.
XX
XX Example 3; Fig 1; 32pp; English.
XX
XX This invention relates to a novel isolated polynucleotide which comprises
XX a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
XX kinase activity and comprising a sequence having 80% identity based on
XX the Clustal alignment method with a fully defined sequence comprising 139
XX amino acids, or its complement. The polynucleotide is useful in genetic
XX or physical mapping for plant breeding. The present sequence is that of a
XX plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
XX exemplification of the invention.
XX
XX Sequence 261 AA;
SQ
XX
XX Query Match 54.3%; Score 44; DB 8; Length 261;
XX Best Local Similarity 63.6%; Pred. NO. 72;
XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 3 LOEFLSHGGYV 13
XX Db 142 LOEFVNHGGVI 152
XX
XX RESULT 45
XX ADJ32018
XX ID ADJ32018 standard; protein; 279 AA.
XX AC ADJ32018;
XX
XX DT 15-APR-2004 (first entry)
XX
XX DE Plant inositol 1,3,4-triphosphate 5/6-kinase amino acid sequence #10.
XX
XX KW inositol 1,3,4-triphosphate 5/6-kinase; genetic mapping;
XX KW physical mapping; plant breeding; plant; enzyme.
XX
XX OS Glycine max.
XX
XX PN US2003229917-A1.
XX
XX PD 11-DEC-2003.
XX
XX PF 20-DEC-2002; 2002US-00324316.
XX
XX PR 11-OCT-2000; 2000US-00696521.
XX
XX PA (CAHO/) CAHOON R E.
XX PA (HITZ/) HITZ W D.
XX PA (PEAR/) PEARLSTEIN R W.
XX PA (CARL/) CARLSON T J.
XX
XX Cahoon RE, Hitz WD, Pearlstein RW, Carlson TJ;
XX WPI; 2004-052054/05.
XX
XX New polynucleotide comprising a sequence encoding a polypeptide having
XX inositol 1,3,4-triphosphate 5/6-kinase activity, useful in genetic or
XX physical mapping for plant breeding.
XX
XX Example 3; Fig 1; 32pp; English.
XX
XX This invention relates to a novel isolated polynucleotide which comprises
XX a sequence encoding a polypeptide having inositol 1,3,4-triphosphate 5/6-
XX kinase activity and comprising a sequence having 80% identity based on
XX the Clustal alignment method with a fully defined sequence comprising 139
XX amino acids, or its complement. The polynucleotide is useful in genetic
XX or physical mapping for plant breeding. The present sequence is that of a
XX plant inositol 1,3,4-triphosphate 5/6-kinase which was used in the
XX exemplification of the invention.
XX
XX Sequence 261 AA;
SQ
XX
XX Query Match 54.3%; Score 44; DB 8; Length 279;
XX Best Local Similarity 63.6%; Pred. NO. 77;
XX Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 3 LOEFLSHGGYV 13
XX Db 173 LOEFVNHGGVI 183
XX
XX RESULT 46
XX AAG44382
XX ID AAG44382 standard; protein; 319 AA.
XX
XX AC AAG44382;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 55589.
XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX
XX OS Arabidopsis thaliana.
XX
XX PN EP1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
```

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 18-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 22-JUN-1999; 99US-0140153P.
PR 23-JUN-1999; 99US-0140354P.
PR 23-JUN-1999; 99US-0140699P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 29-JUN-1999; 99US-0141287P.
PR 30-JUN-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0146703P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 319;

Best Local Similarity 63.6%; Pred. No. 88;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13

||||:||||:

Db 176 LOEFVNHGGVI 186

RESULT 47

AAV59421
 ID AAV59421 standard; protein; 321 AA.

XX AC AAV59421;

XX DT 21-MAR-2000 (first entry)

XX DE Corn inositol 1,3,4-triphosphate 5/6-kinase protein sequence.

XX KW Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
 KW myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
 KW animal feed.

XX OS Zea mays.

XX PN WO9955879-A1.

XX PD 04-NOV-1999.

XX PF 22-APR-1999; 99WO-US008790.

XX PR 24-APR-1998; 98US-0082960P.

XX XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;

XX PS WPI; 2000-072179/06.

XX DR N-PSDB; AAZ48805.

XX PT Novel phytic acid biosynthetic enzymes used to alter the level of the
 XX enzyme in transgenic plants.

XX PS Claim 5; Page 44-45; 63pp; English.

XX CC This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein
 CC of the invention. ITK is a phytic acid [Myo-inositol 1,2,3,4,5,6-
 CC hexaphosphate] biosynthetic enzyme. The ITK enzymes of the invention may
 CC be prepared recombinantly and used to raise antibodies, which are used
 CC for detecting the enzymes in situ in cells or in vitro in cell extracts.
 CC The polynucleotides of the invention may be used to create transgenic
 CC plants in which the ITK levels are present at higher or lower levels than
 CC normal, or in cell types or developmental processes where they are not
 CC normally found. This would alter the level of 1,3,4-triphosphate 5/6
 CC kinase found in those cells. In addition, it may be desirable to
 CC eliminate expression of ITK genes for certain applications. The
 CC polynucleotides also serve as a source of probes and primers, which are
 CC useful for genetic mapping, as markers for traits linked to those genes,
 CC and to isolate homologous sequences from other species, as well as for
 CC physical mapping. Decreased amounts of phytate and increased amounts of
 CC available phosphate in animal feeds such as corn would lead to improved
 CC feed efficiency. The proteins of the present invention lead to a better
 CC understanding of the phytic acid biosynthesis pathway, allowing it to be
 CC exploited for commercial uses, e.g. in animal feeds

XX SQ Sequence 321 AA;

Query Match 54.3%; Score 44; DB 3; Length 321;

Best Local Similarity 63.6%; Pred. No. 89;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13

||||:||||:

Db 185 LOEFVNHGGVI 195

RESULT 48

AAV59423
 ID AAV59423 standard; protein; 341 AA.

XX AC AAV59423;

XX DT 21-MAR-2000 (first entry)

XX DE Rice inositol 1,3,4-triphosphate 5/6-kinase protein sequence.

XX KW Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
 KW myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
 KW animal feed.

XX OS Oryza sativa.

XX PN WO9955879-A1.

XX PD 04-NOV-1999.

XX PF 22-APR-1999; 99WO-US008790.

XX PR 24-APR-1998; 98US-0082960P.

XX XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;

XX PS WPI; 2000-072179/06.

XX DR N-PSDB; AAZ48807.

XX PT Novel phytic acid biosynthetic enzymes used to alter the level of the
 XX enzyme in transgenic plants.

XX PS Claim 5; Page 47-48; 63pp; English.

XX CC This sequence is an inositol 1,3,4-triphosphate 5/6-kinase (ITK) protein
 CC of the invention. ITK is a phytic acid [Myo-inositol 1,2,3,4,5,6-
 CC hexaphosphate] biosynthetic enzyme. The ITK enzymes of the invention may
 CC be prepared recombinantly and used to raise antibodies, which are used
 CC for detecting the enzymes in situ in cells or in vitro in cell extracts.
 CC The polynucleotides of the invention may be used to create transgenic
 CC plants in which the ITK levels are present at higher or lower levels than
 CC normal, or in cell types or developmental processes where they are not
 CC normally found. This would alter the level of 1,3,4-triphosphate 5/6
 CC kinase found in those cells. In addition, it may be desirable to
 CC eliminate expression of ITK genes for certain applications. The
 CC polynucleotides also serve as a source of probes and primers, which are
 CC useful for genetic mapping, as markers for traits linked to those genes,
 CC and to isolate homologous sequences from other species, as well as for
 CC physical mapping. Decreased amounts of phytate and increased amounts of
 CC available phosphate in animal feeds such as corn would lead to improved
 CC feed efficiency. The proteins of the present invention lead to a better
 CC understanding of the phytic acid biosynthesis pathway, allowing it to be
 CC exploited for commercial uses, e.g. in animal feeds

XX SQ Sequence 341 AA;

Query Match 54.3%; Score 44; DB 3; Length 341;

Best Local Similarity 63.6%; Pred. No. 94;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13

||||:||||:

Db 185 LOEFVNHGGVI 195


```
XX
SQ Sequence 352 AA;
  Query Match      54.3%; Score 44; DB 8; Length 352;
  Best Local Similarity 63.6%; Pred. No. 97;
  Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      3 LQEFLSHGGYV 13
      |||:|:|:|
Db     194 LQEFVNHGGVI 204

Search completed: May 12, 2006, 10:37:20
Job time : 130.809 secs
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16,9915 Seconds

(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-53

Perfect score: 81

Sequence: 1 CSLQFLSHGGYVC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR 80.*

1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	59.3	660	2 A32378	probable ATP-depen
2	48	59.3	662	1 I84741	RNA helicase - mou
3	45	55.6	536	2 A81710	ribosomal protein
4	45	55.6	697	1 S13654	ATP-dependent RNA
5	44	54.3	319	2 JC5401	inositol 1,3,4-tri
6	44	54.3	429	2 A38220	transforming prote
7	42	51.9	293	2 T19954	hypothetical prote
8	42	51.9	338	2 T10544	inositol 1,3,4-tri
9	42	51.9	439	2 B72313	glutamine syntheta
10	41.5	51.2	113	2 H90876	probable minor tai
11	41	50.6	436	2 T23345	hypothetical prote
12	41	50.6	470	2 I38026	MLN 62 protein - h
13	40.5	50.0	880	2 G88925	protein T22H9.2 [i
14	40	49.4	392	2 PQ0446	DNA-binding protein
15	40	49.4	461	2 A46394	suppressor protein
16	40	49.4	497	2 A83884	L-arabinose isomer
17	40	49.4	569	2 A71556	probable S1 riboso
18	40	49.4	643	2 T27429	hypothetical prote
19	40	49.4	772	2 D56695	transducin-like en
20	39.5	48.8	113	2 D90854	minor tail protein
21	39.5	48.8	113	2 B90823	probable minor tai
22	39.5	48.8	161	2 C95858	probable oxidoredu
23	39	48.1	187	2 B75576	oxidoreductase, ir
24	39	48.1	187	2 B75576	oxidoreductase, ir
25	39	48.1	213	2 F71804	probable transcript
26	39	48.1	525	2 T21527	hypothetical prote
27	39	48.1	532	2 C90154	primase SS00140 [i
28	39	48.1	646	2 T35002	probable respirato
29	39	48.1	652	2 B83725	hypothetical prote

30	39	48.1	891	2 AG0265	alcohol dehydrogen
31	39	48.1	892	2 AE0650	alcohol dehydrogen
32	39	48.1	1049	1 S19421	ATP-dependent perm
33	39	48.1	1588	1 BVBYA1	3-dehydroquinatase
34	39	48.1	1642	2 T19130	hypothetical prote
35	38.5	47.5	1878	2 E86189	hypothetical prote
36	38	46.9	94	2 T03751	hypothetical prote
37	38	46.9	105	2 A82985	hypothetical prote
38	38	46.9	142	2 S46445	lectin precursor -
39	38	46.9	160	2 JE0136	response regulator
40	38	46.9	213	2 E64690	endonuclease III X
41	38	46.9	218	2 E82780	endonuclease III -
42	38	46.9	259	2 D75275	transcription regu
43	38	46.9	295	2 AH2340	starch synthase DU
44	38	46.9	310	2 T01266	phosphonates trans
45	38	46.9	323	2 D83708	hypothetical prote
46	38	46.9	421	2 E71363	probable glucose-6
47	38	46.9	425	2 E64657	Glutamate dehydrog
48	38	46.9	447	1 A33504	NADP-specific glut
49	38	46.9	447	2 AF0710	conserved hypother
50	38	46.9	471	2 E83130	hypothetical prote
51	38	46.9	670	2 AH2425	ABC1-type transpor
52	38	46.9	681	2 S76354	hypothetical prote
53	38	46.9	961	2 E86245	chromosome partiti
54	37.5	46.3	262	2 B82950	tMDC II protein -
55	37.5	46.3	756	2 S47656	hypothetical prote
56	37	45.7	101	2 C72537	hypothetical prote
57	37	45.7	177	2 S77911	DNA-(apurinic or a
58	37	45.7	210	2 E84943	oxidoreductase, ir
59	37	45.7	214	2 D75576	3-methyl-2-oxobuta
60	37	45.7	246	2 E82832	conserved hypother
61	37	45.7	255	2 G87677	hypothetical prote
62	37	45.7	314	2 T28879	hypothetical prote
63	37	45.7	416	2 T26371	hypothetical prote
64	37	45.7	421	2 T23981	glutamate-ammonia
65	37	45.7	469	2 S23899	2-hydroxyacylasp
66	37	45.7	541	2 A48801	hypothetical prote
67	37	45.7	716	2 T37710	CWH41 protein - ye
68	37	45.7	833	2 S62136	chitin synthase (E
69	37	45.7	1026	2 T18220	hypothetical prote
70	37	45.7	1142	2 T30272	chromosome-associ
71	37	45.7	1244	2 T23744	hypothetical prote
72	37	45.7	1620	2 T27283	rad3 checkpoint pr
73	37	45.7	2386	2 T39911	notch protein homo
74	37	45.7	2531	2 S18188	neurofibromatosis-
75	37	45.7	2818	2 B55282	neurofibromin I -
76	37	45.7	2820	2 JC5196	neurofibromin I -
77	37	45.7	2825	2 I54352	probable protein k
78	36.5	45.1	379	2 T47507	coat protein - tob
79	36	44.4	156	1 VCTMHR	regulatory protein
80	36	44.4	223	2 T44784	Ig kappa chain - s
81	36	44.4	230	2 S33161	flagellar L-ring p
82	36	44.4	238	2 H84969	hypothetical prote
83	36	44.4	245	2 T51451	probable bacteriop
84	36	44.4	249	2 AF0735	thiamin biosynthes
85	36	44.4	267	2 A64375	kanamycin kinase (
86	36	44.4	271	2 JU0103	hypothetical prote
87	36	44.4	282	2 AF0332	28k surface antige
88	36	44.4	287	2 JE0220	interleukin 12 p40
89	36	44.4	335	2 I72789	leucine dehydrogen
90	36	44.4	351	2 F72020	Leucine dehydrogen
91	36	44.4	351	2 E86605	probable secreted
92	36	44.4	360	2 T35783	hypothetical prote
93	36	44.4	364	2 T04994	probable periplasm
94	36	44.4	397	2 B81839	NADH2 dehydrogenas
95	36	44.4	426	1 E70351	NADH2 dehydrogenas
96	36	44.4	438	2 T11903	UDP-glucuronosyltr
97	36	44.4	508	2 T03910	cell polarity prot
98	36	44.4	522	2 T37536	hypothetical prote
99	36	44.4	531	2 T23835	hypothetical prote
100	36	44.4	535	2 T24309	pectate disacchari
101	36	44.4	555	2 AC0210	hypothetical prote
102	36	44.4	599	2 T16774	hypothetical prote

103	36	44.4	601	1	S11777	hydrogenase (EC 1.	176	35	43.2	740	2	B86638	GTP diphosphokinas
104	36	44.4	602	1	JQ0806	hydrogenase (EC 1.	177	35	43.2	860	1	QRHULD	LDL receptor precu
105	36	44.4	679	2	R83723	hypothetical prote	178	35	43.2	911	2	JC6016	chitin synthase (E
106	36	44.4	764	2	D84847	probable receptor-	179	35	43.2	916	2	JC2315	chitin synthase (E
107	36	44.4	780	2	D96803	probable beta-gala	180	35	43.2	916	2	A81089	DNA gyrase chain A
108	36	44.4	803	2	H87197	penicillin binding	181	35	43.2	916	2	B81853	DNA topoisomerase
109	36	44.4	834	2	T29821	hypothetical prote	182	35	43.2	916	2	S60779	DNA gyrase chain A
110	36	44.4	839	2	T45908	hypothetical prote	183	35	43.2	916	2	JT0396	reverse transcript
111	36	44.4	891	1	DBEC	acetaldehyde dehyd	184	35	43.2	943	2	JE0121	hypothetical 107.4
112	36	44.4	891	2	D85704	hypothetical prote	185	35	43.2	957	2	E86489	protein T32820.19
113	36	44.4	891	2	E90846	acetaldehyde dehyd	186	35	43.2	961	2	AD0548	putative autotransp
114	36	44.4	928	1	VGBBEG	acetaldehyde dehyd	187	35	43.2	996	2	A71080	hypothetical prote
115	36	44.4	932	1	VGBBEC	glycoprotein GI pr	188	35	43.2	1184	2	A96638	hypothetical prote
116	36	44.4	938	2	AB1530	transcription regu	189	35	43.2	1206	2	D90085	hypothetical prote
117	36	44.4	938	2	AL1172	transcription regu	190	35	43.2	1899	2	T49273	hypothetical prote
118	36	44.4	1056	2	H84777	probable kinesin-r	191	35	43.2	1984	2	T13171	probable vielloge
119	36	44.4	1059	2	T21891	hypothetical prote	192	35	43.2	2204	1	RRNZNV	genome polyproteic
120	36	44.4	1156	2	T14899	aminophospholipid	193	35	43.2	4543	1	A53102	alpha-2-macroglubu
121	35.5	43.8	334	2	T03764	protein RWD - rice	194	34.5	42.6	347	1	C46137	opsin, violet-sens
122	35.5	43.8	363	2	AC1700	peptidoglycan synt	195	34.5	42.6	363	2	E83970	UDP-N-acetylglucos
123	35.5	43.8	609	2	F84824	hypothetical prote	196	34.5	42.6	524	2	S38539	disintegrin-like m
124	35.5	43.8	682	2	T12968	hypothetical prote	197	34.5	42.6	670	2	I65967	disintegrin-like m
125	35.5	43.8	1097	1	S47220	protein kinase C (198	34.5	42.6	720	2	A36526	choline O-acetyltr
126	35.5	43.8	1148	2	T18770	probable calcium c	199	34.5	42.6	728	2	A24889	subtilisin-like pr
127	35.5	43.8	1412	2	T01610	RNA-directed DNA p	200	34.5	42.6	915	1	A48225	choline O-acetyltr
128	35	43.2	83	2	A82782	hypothetical prote	201	34.5	42.6	915	2	B48225	probable proprotel
129	35	43.2	103	2	AD1276	hypothetical prote	202	34.5	42.6	1548	2	S34583	serine proteinase
130	35	43.2	103	2	AD1639	hypothetical prote	203	34	42.0	40	2	F84316	hypothetical prote
131	35	43.2	103	2	AD2089	hypothetical prote	204	34	42.0	68	2	H82124	heme exporter prot
132	35	43.2	174	2	G90432	conserved hypotet	205	34	42.0	98	2	AG2307	hypothetical prote
133	35	43.2	187	2	A87512	conserved hypotet	206	34	42.0	139	2	S67187	hypothetical prote
134	35	43.2	210	2	B54639	oxygen-insensitive	207	34	42.0	148	2	B82387	conserved hypotet
135	35	43.2	222	2	T30423	hypothetical prote	208	34	42.0	151	2	T32957	hypothetical prote
136	35	43.2	234	2	H85613	leucyl, phenylalan	209	34	42.0	152	2	F71973	hypothetical prote
137	35	43.2	234	2	A36898	leucyl, phenylalan	210	34	42.0	163	2	S43764	mannose-binding le
138	35	43.2	234	2	B90750	leucyl, phenylalan	211	34	42.0	163	2	S43763	mannose-binding le
139	35	43.2	237	1	S56137	membrane-bound tet	212	34	42.0	171	2	T49538	hypothetical prote
140	35	43.2	249	2	T04436	ankyrin 3 homolog	213	34	42.0	186	2	D64533	hypothetical prote
141	35	43.2	256	2	F70525	probable gp92 pro	214	34	42.0	200	2	H81369	probable 3-isoprop
142	35	43.2	257	2	AB0353	acetyltransferase	215	34	42.0	221	2	B81781	beta-phosphoglucos
143	35	43.2	269	2	A48906	nicotinate-nucleot	216	34	42.0	221	2	G81203	beta-phosphoglucos
144	35	43.2	273	2	H71827	conserved hypotet	217	34	42.0	221	2	F81780	probable anti-sigm
145	35	43.2	290	2	T46470	hypothetical prote	218	34	42.0	226	2	F82091	hypothetical prote
146	35	43.2	303	2	A72365	conserved hypotet	219	34	42.0	226	2	A70436	hypothetical prote
147	35	43.2	312	2	C84030	3-oxoacyl-(acyl-ca	220	34	42.0	242	2	B86293	F7H2.20 protein -
148	35	43.2	320	2	AC2315	hypothetical prote	221	34	42.0	254	2	H72102	3-deoxy-manno-octu
149	35	43.2	330	2	AD2082	iron(III) dicitrat	222	34	42.0	254	2	C86520	deoxyoctulosonic a
150	35	43.2	344	2	F72406	hypothetical prote	223	34	42.0	255	2	S74930	hypothetical prote
151	35	43.2	367	1	D70395	conserved hypotet	224	34	42.0	256	2	T15383	hypothetical prote
152	35	43.2	383	2	T34718	probable integral	225	34	42.0	262	2	B90450	conserved hypotet
153	35	43.2	390	2	UC4023	transforming growt	226	34	42.0	263	1	SNRTC2	proteasome endopep
154	35	43.2	396	2	B83461	hypothetical prote	227	34	42.0	263	2	A70163	hemolysin (tlyA) h
155	35	43.2	400	2	C97147	probable glycosylt	228	34	42.0	264	1	WMBVT3	30K protein - toma
156	35	43.2	425	2	A83358	phosphopyruvate hy	229	34	42.0	268	2	AG0332	conserved hypotet
157	35	43.2	428	2	AF1984	hypothetical prote	230	34	42.0	269	1	JC1445	proteasome endopep
158	35	43.2	430	2	H69364	aspartyl-tRNA synt	231	34	42.0	271	1	F69656	naphthoate synthas
159	35	43.2	440	2	G87444	8-amino-7-oxononan	232	34	42.0	272	2	AL1283	dihydroxynaphthoic
160	35	43.2	444	1	Q4ADA7	probable thiophene	233	34	42.0	272	2	AD1655	dihydroxynaphthoic
161	35	43.2	448	1	Q4ADA7	maturation (pIva2)	234	34	42.0	274	2	F75328	conserved hypotet
162	35	43.2	449	1	Q4ADA2	maturation (pIva2)	235	34	42.0	275	2	B70640	hypothetical prote
163	35	43.2	449	1	Q4ADA5	maturation (pIva2)	236	34	42.0	275	2	G83107	probable two-compo
164	35	43.2	452	2	S33932	maturation protein	237	34	42.0	277	2	A97080	pns system, fructo
165	35	43.2	535	2	C95057	CTP synthase [lmpo	238	34	42.0	288	2	B81650	endonuclease IV TC
166	35	43.2	535	2	F97926	CTP synthase (BC 6	239	34	42.0	288	2	B71491	probable endonucle
167	35	43.2	544	2	S25101	CTP synthase (EC 6	240	34	42.0	290	2	B71152	probable drdp-4-de
168	35	43.2	544	2	I55454	neuroglycan C prec	241	34	42.0	290	2	AF1133	hypothetical prote
169	35	43.2	548	2	T16642	hypothetical prote	242	34	42.0	301	1	LNRT2	hepatic lectin 2 -
170	35	43.2	550	2	H64205	phosphomannomutase	243	34	42.0	310	2	SA5130	autophagocytosis p
171	35	43.2	605	2	F84109	two-component sens	244	34	42.0	314	2	I38964	E18.19K/Bcl-2-inte
172	35	43.2	618	1	B43255	hydrogenase (EC 1.	245	34	42.0	318	2	C18190	probable sodium-tr
173	35	43.2	626	2	T49625	glucan 1.4-alpha-g	246	34	42.0	325	1	A45470	hydroxymethylgluta
174	35	43.2	626	2	S36364	glucan 1.4-alpha-g	247	34	42.0	325	1	A71887	probable GMP reduc
175	35	43.2	708	2	AH0192	alpha-galactosidas	248	34	42.0	327	1	F64626	probable GMP reduc

249	34	42.0	333	1	B53308	mosa protein - Rhi	322	33.5	41.4	2083	2	T42721	CRP-ductin-alpha p
250	34	42.0	335	2	C64380	porphobilinogen sy	323	33.5	41.4	2524	2	A35844	Xotch protein - Af
251	34	42.0	350	2	H96768	protein flavonol s	324	33	40.7	49	2	F82535	hypothetical prote
252	34	42.0	361	2	A87132	probable alcohol d	325	33	40.7	54	2	AH1965	hypothetical prote
253	34	42.0	361	2	G70862	probable adh2 pro	326	33	40.7	95	2	H71256	conserved hypotet
254	34	42.0	364	2	B36313	hypothetical 42K p	327	33	40.7	98	2	T42909	hypothetical prote
255	34	42.0	370	2	S70157	cpaf protein, 40,6	328	33	40.7	104	2	F53275	Ig kappa-1 chain C
256	34	42.0	370	2	T16213	APX-1 protein homo	329	33	40.7	131	2	I56894	complement regulat
257	34	42.0	405	2	A34851	lariat-debranching	330	33	40.7	132	1	HAFGH	hemoglobin alpha-t
258	34	42.0	406	2	A54857	transcription fact	331	33	40.7	133	2	AB3465	complement regulat
259	34	42.0	406	2	A19655	hypothetical prote	332	33	40.7	147	2	A69933	hypothetical prote
260	34	42.0	420	2	T00629	hypothetical prote	333	33	40.7	154	2	A89962	hypothetical prote
261	34	42.0	457	2	AF2819	amine oxidase, fla	334	33	40.7	162	2	S43761	mannose-binding le
262	34	42.0	457	2	F97597	glycerol-3-phospha	335	33	40.7	173	2	S77232	hypothetical prote
263	34	42.0	473	2	AC0479	hypothetical prote	336	33	40.7	177	2	E87240	probable peptide m
264	34	42.0	473	2	T46999	hypothetical prote	337	33	40.7	193	2	A10387	hypothetical prote
265	34	42.0	473	2	AC0239	probable GntR-fam1	338	33	40.7	196	2	A99170	conserved hypotet
266	34	42.0	483	2	H85073	probable transposo	339	33	40.7	196	2	D87712	imidazoleglycerol-
267	34	42.0	485	2	T01244	hypothetical prote	340	33	40.7	208	2	F84942	riboflavin synthas
268	34	42.0	488	2	T45964	LAX1 / AUX1-like p	341	33	40.7	211	2	T41710	hypothetical prote
269	34	42.0	501	2	T31754	hypothetical prote	342	33	40.7	211	2	S77532	rehydrin - Synecho
270	34	42.0	504	2	T34106	hypothetical prote	343	33	40.7	214	2	G90582	adenylate kinase (
271	34	42.0	504	2	C64398	hypothetical prote	344	33	40.7	214	2	A87078	probable 3-methylp
272	34	42.0	510	2	T33770	hypothetical prote	345	33	40.7	216	1	QBBER8	9R protein - human
273	34	42.0	510	2	T48977	carboxypeptidase-1	346	33	40.7	218	2	T44209	hypothetical prote
274	34	42.0	514	2	T25534	hypothetical prote	347	33	40.7	218	2	T44023	hypothetical prote
275	34	42.0	525	2	T24647	hypothetical prote	348	33	40.7	219	1	H64880	probable beta-phos
276	34	42.0	532	2	D63281	CTP synthase (pyrG	349	33	40.7	219	2	A85753	probable beta-phos
277	34	42.0	535	2	G97255	CTP synthase (UTP-	350	33	40.7	219	2	H90865	hypothetical prote
278	34	42.0	553	2	G71543	probable fructose-	351	33	40.7	235	2	E36811	probable beta-phos
279	34	42.0	560	2	S27874	steroid hormone re	352	33	40.7	235	2	A75293	probable beta-phos
280	34	42.0	578	2	D79555	CTP synthase (UTP-	353	33	40.7	243	2	T46221	hypothetical prote
281	34	42.0	578	2	AP2775	CTP synthase (impo	354	33	40.7	250	2	H84962	hypothetical prote
282	34	42.0	594	1	E64622	excinuclease ABC c	355	33	40.7	252	2	C82507	hypothetical prote
283	34	42.0	594	2	B71893	excinuclease ABC c	356	33	40.7	253	2	T35768	hypothetical prote
284	34	42.0	627	2	B44009	gamma-aminobutyric	357	33	40.7	258	2	D71280	hypothetical prote
285	34	42.0	627	2	JH0695	gamma-aminobutyric	358	33	40.7	262	2	I49361	natural killer cel
286	34	42.0	665	2	B96668	probable acyl-CoA	359	33	40.7	262	2	A45813	T-cell surface gly
287	34	42.0	693	2	T13152	hypothetical prote	360	33	40.7	262	2	G90018	conserved hypotet
288	34	42.0	704	2	B84685	hypothetical prote	361	33	40.7	268	2	T38024	hypothetical prote
289	34	42.0	722	2	B71376	hypothetical prote	362	33	40.7	278	2	T38024	conserved hypotet
290	34	42.0	745	2	T23893	hypothetical prote	363	33	40.7	282	2	F82851	hypothetical prote
291	34	42.0	764	2	S48521	AKR1 protein - yea	364	33	40.7	290	2	AH3448	hypothetical prote
292	34	42.0	767	2	G96661	hypothetical prote	365	33	40.7	292	2	D71285	conserved hypotet
293	34	42.0	775	1	VPXRFPG	outer layer protei	366	33	40.7	292	2	T15973	hypothetical prote
294	34	42.0	820	2	G88996	protein C17B7.5 [i	367	33	40.7	304	2	T12891	hypothetical prote
295	34	42.0	833	2	AB0448	probable insectici	368	33	40.7	313	1	H69510	conserved hypotet
296	34	42.0	841	2	T01011	hypothetical prote	369	33	40.7	320	2	E95315	hypothetical prote
297	34	42.0	972	2	A30363	glycoprotein Gp330	370	33	40.7	322	2	AH1344	hypothetical prote
298	34	42.0	1009	2	T32464	hypothetical prote	371	33	40.7	322	2	AC1715	hypothetical prote
299	34	42.0	1184	2	T09484	cartilage intermed	372	33	40.7	325	2	I57009	hypothetical prote
300	34	42.0	1245	2	D71613	GAF domain protein	373	33	40.7	326	2	T22407	hypothetical prote
301	34	42.0	1458	2	S36014	dynamin heavy chain	374	33	40.7	327	2	F82038	hypothetical prote
302	34	42.0	1463	2	C95032	hypothetical prote	375	33	40.7	334	2	B44478	probable cell grow
303	34	42.0	1463	2	C97903	DNA-directed DNA p	376	33	40.7	337	2	S55657	adenosine deaminas
304	34	42.0	1650	2	S53457	dominant autoantig	377	33	40.7	337	2	B44478	probable cell grow
305	34	42.0	1650	2	A46019	notch-1 protein -	378	33	40.7	338	2	S55657	capacid protein 62
306	34	42.0	2531	2	A46019	notch protein homo	379	33	40.7	349	2	S22448	FMRI protein - Pod
307	34	42.0	2555	2	A40043	inositol 1,4,5-tri	380	33	40.7	349	2	S22448	DNA/RNA-binding pr
308	34	42.0	2670	2	A46719	inositol 1,4,5-tri	381	33	40.7	363	2	I38937	5S rRNA-binding pr
309	34	42.0	2671	2	A49873	inositol 1,4,5-tri	382	33	40.7	365	2	A34895	lignin peroxidase
310	34	42.0	2788	2	T30851	lysosomal traffick	383	33	40.7	367	2	S47649	hypothetical prote
311	34	42.0	3788	2	T13360	lysosomal traffick	384	33	40.7	368	2	F98282	histidinol-phospha
312	34	42.0	3796	2	T18514	alpha-2-macroglobu	385	33	40.7	368	2	AB3001	probable cell grow
313	34	42.0	4544	1	S02392	gp330 protein prec	386	33	40.7	373	2	A44478	protein-tyrosine-p
314	34	42.0	4660	2	T42737	membrane-bound imm	387	33	40.7	377	1	A48711	conserved hypotet
315	34	42.0	4660	2	A46477	hypothetical prote	388	33	40.7	378	2	D69793	conserved hypotet
316	34	42.0	307	2	C70734	phospho-N-acetylmu	389	33	40.7	379	2	E81081	hypothetical prote
317	34	42.0	363	2	JC1275	hypothetical prote	390	33	40.7	391	2	F81861	G-box-binding prot
318	34	42.0	376	2	T16116	hypothetical prote	391	33	40.7	406	2	C64998	probable sulfatase
319	34	42.0	429	2	F84015	maltose/maltodextr	392	33	40.7	407	2	F85867	probable sulfatase
320	34	42.0	438	2	C86244	DnaJ homolog, 4706	393	33	40.7	416	2	E91023	probable sarcosine
321	34	42.0	849	2	H84668	Mutator-like trans	394	33	40.7	416	2	D84638	

395	33	40.7	425	2	H70793	hypothetical prote	468	33	40.7	1026	2	A49750	beta-galactosidase
396	33	40.7	427	2	P75406	histidyl-tRNA synt	469	33	40.7	1054	2	T14189	hypothetical prote
397	33	40.7	431	2	A80410	phosphopyruvate hy	470	33	40.7	1133	1	EGRT	epidermal growth f
398	33	40.7	432	1	NOEC	phosphopyruvate hy	471	33	40.7	1133	1	T12529	hypothetical prote
399	33	40.7	432	2	H85928	enolase [imported]	472	33	40.7	1207	1	EGHU	epidermal growth f
400	33	40.7	432	2	G91083	enolase [imported]	473	33	40.7	1217	1	EGMSG	epidermal growth f
401	33	40.7	432	2	AC0859	enolase [imported]	474	33	40.7	1274	2	I40813	neurotoxin type F
402	33	40.7	432	2	G90268	conserved hypothet	475	33	40.7	1304	2	T14073	neutrin lb heavy ch
403	33	40.7	433	2	B81992	NADH2 dehydrogenas	476	33	40.7	1346	2	T14849	cytosadherence link
404	33	40.7	433	2	D81222	NADH dehydrogenase	477	33	40.7	1449	2	T20181	hypothetical prote
405	33	40.7	433	2	C84619	probable serine ca	478	33	40.7	1514	2	S70099	hypothetical prote
406	33	40.7	439	1	ISHIX	xylose isomerase (479	33	40.7	1518	2	PQ0221	polyprotein - plum
407	33	40.7	442	2	C81370	probable thiophene	480	33	40.7	1573	2	AF3514	glutamate synthase
408	33	40.7	456	2	E69391	hypothetical prote	481	33	40.7	1583	2	P86366	protein P26F24.8 l
409	33	40.7	459	2	T51558	probable flavonol	482	33	40.7	1584	2	T20180	hypothetical prote
410	33	40.7	460	2	T04579	hypothetical prote	483	33	40.7	1586	2	T20179	hypothetical prote
411	33	40.7	464	2	T47710	glucuronosyl trans	484	33	40.7	1597	2	S65053	genome polyprotein
412	33	40.7	465	2	D96567	F6D8.13 [imported]	485	33	40.7	1699	2	T14074	complement compone
413	33	40.7	468	2	H70201	UDP-N-acetylmurama	486	33	40.7	2264	1	GNVVTB	genome polyprotein
414	33	40.7	469	2	B37837	probable alpha-amy	487	33	40.7	2276	2	T00076	hypothetical prote
415	33	40.7	472	2	T51559	probable flavonol	488	33	40.7	2330	1	RRIMWV	genome polyprotein
416	33	40.7	479	2	B69784	transcription regu	489	33	40.7	2962	2	T19756	hypothetical prote
417	33	40.7	485	2	S36772	E-selectin - bovin	490	33	40.7	3131	2	S39842	emiatin synthetas
418	33	40.7	493	2	S32037	finger protein XFG	491	33	40.7	3225	2	D81702	adherence factor T
419	33	40.7	497	2	T37827	hypothetical prote	492	33	40.7	3635	2	T10053	laminin alpha 5 ch
420	33	40.7	506	2	C81704	monooxygenase-rela	493	33	40.7	3973	2	B71612	hypothetical prote
421	33	40.7	512	2	T40576	probable structure	494	32.5	40.1	153	2	PN0103	hypothetical lTK p
422	33	40.7	519	2	UC4762	RNA-directed RNA p	495	32.5	40.1	196	2	AF3358	methyltransferase
423	33	40.7	526	2	T25535	hypothetical prote	496	32.5	40.1	258	2	AB1956	hypothetical prote
424	33	40.7	527	2	I84483	protein-tyrosine k	497	32.5	40.1	316	2	S57839	CPC2 protein - Neu
425	33	40.7	533	1	S75536	NADH2 dehydrogenas	498	32.5	40.1	363	2	AC1329	peptidoglycan synt
426	33	40.7	533	2	T15573	hypothetical prote	499	32.5	40.1	409	2	B37753	Nadr protein - Sal
427	33	40.7	541	2	UC5443	2-hydroxacylsphin	500	32.5	40.1	476	2	H83265	probable D-alanyl-
428	33	40.7	550	1	A34576	crystal protein pr	501	32.5	40.1	520	2	B71443	hypothetical prote
429	33	40.7	553	2	C81698	pyrophosphate-fruc	502	32.5	40.1	2109	2	I38414	transcription fact
430	33	40.7	567	2	T25771	hypothetical prote	503	32.5	40.1	2946	2	T15840	hypothetical prote
431	33	40.7	596	1	S11969	hydrogenase (EC 1.	504	32	39.5	53	2	AG0014	hypothetical prote
432	33	40.7	598	2	T42592	protein-serine/thr	505	32	39.5	59	2	G81250	probable transcrit
433	33	40.7	599	2	B73268	conserved hypothet	506	32	39.5	63	2	T41452	probable protein t
434	33	40.7	610	2	A35046	E-selectin precurs	507	32	39.5	65	2	C82542	hypothetical prote
435	33	40.7	628	2	A87596	hypothetical prote	508	32	39.5	80	2	AH2958	hypothetical prote
436	33	40.7	633	2	C70358	hydrogenase (EC 1.	509	32	39.5	80	2	F98324	hypothetical prote
437	33	40.7	634	2	S33339	somatotropin recep	510	32	39.5	87	2	T03315	gene l8 protein -
438	33	40.7	635	1	T37835	probable phosphoe	511	32	39.5	91	2	JE0321	ribosomal protein
439	33	40.7	638	2	A33991	somatotropin recep	512	32	39.5	91	2	C83414	hypothetical prote
440	33	40.7	638	2	S12136	somatotropin recep	513	32	39.5	102	2	T32603	hypothetical prote
441	33	40.7	638	2	B28176	somatotropin recep	514	32	39.5	116	2	D81945	probable pilus bio
442	33	40.7	640	2	S62747	homeotic protein A	515	32	39.5	116	2	G81159	probable type IV p
443	33	40.7	642	2	B69144	probable formate C	516	32	39.5	122	2	B23843	serum amyloid prot
444	33	40.7	653	2	T51895	hypothetical prote	517	32	39.5	122	2	A23843	serum amyloid prot
445	33	40.7	662	2	G87630	methylmalonyl-CoA	518	32	39.5	122	2	I49496	anlyoid A - mouse
446	33	40.7	665	2	T18979	hypothetical prote	519	32	39.5	127	2	F83803	hypothetical prote
447	33	40.7	712	1	A47331	ribonucleoside-tri	520	32	39.5	132	2	T26527	two-component resp
448	33	40.7	712	2	AC1058	ribonucleoside-tri	521	32	39.5	140	2	AG2146	hypothetical prote
449	33	40.7	712	2	AG0419	ribonucleoside-tri	522	32	39.5	140	2	AC1264	hypothetical prote
450	33	40.7	712	2	G91280	anaerobic ribonuc	523	32	39.5	140	2	AE1626	hypothetical prote
451	33	40.7	712	2	G86121	anaerobic ribonuc	524	32	39.5	149	2	AB1395	hypothetical prote
452	33	40.7	721	2	T08956	AIM1 protein - Ara	525	32	39.5	149	2	AE1770	hypothetical prote
453	33	40.7	739	2	A83015	primosomal protein	526	32	39.5	157	2	S19735	lectin precursor -
454	33	40.7	775	2	AH2803	phosphoribosylform	527	32	39.5	157	2	T40106	conserved hypothet
455	33	40.7	775	2	H97592	phosphoribosylform	528	32	39.5	157	2	F87509	conserved hypothet
456	33	40.7	790	1	S77032	ABC transporter sl	529	32	39.5	163	2	B84731	hypothetical prote
457	33	40.7	830	2	A84951	DNA topoisomerase	530	32	39.5	163	2	T10220	hypothetical prote
458	33	40.7	868	1	VGBE31	Glycoprotein B - h	531	32	39.5	167	2	B83743	4-hydroxybenzoyl-C
459	33	40.7	872	2	S73785	DNA polymerase III	532	32	39.5	173	2	A70017	4-hydroxybenzoyl-C
460	33	40.7	879	1	QRRTLD	LDL receptor precu	533	32	39.5	180	2	F82187	TonB system transp
461	33	40.7	923	2	H83248	DNA gyrase subunit	534	32	39.5	183	1	C70008	pyrazinamidase/nic
462	33	40.7	942	2	C83861	ATP-dependent DNA	535	32	39.5	184	2	S77751	probable ABC-type
463	33	40.7	946	1	A87239	ror-related recept	536	32	39.5	190	2	F87412	conserved hypothet
464	33	40.7	953	2	T40643	probable serine th	537	32	39.5	192	2	H82467	hypothetical prote
465	33	40.7	1020	2	T51925	protein envelope c	538	32	39.5	193	2	D81006	conserved hypothet
466	33	40.7	1020	2	T51925	Ca2+-transporting	539	32	39.5	195	2	B27733	nifQ protein - Azo
467	33	40.7	1020	2	T51926	Ca2+-transporting	540	32	39.5	196	2	D69275	conserved hypothet

541	32	39.5	200	2	I48615	gene KIS protein -	614	32	39.5	357	2	T05625	cinnamyl-alcohol d
542	32	39.5	203	2	G70501	probable 3-methylp	615	32	39.5	358	2	F98289	lps biosynthesis r
543	32	39.5	204	2	B75539	probable phosphogl	616	32	39.5	359	2	F90055	conserved hypotnet
544	32	39.5	212	2	H83208	endonuclease III P	617	32	39.5	362	2	S33733	G protein-coupled
545	32	39.5	218	2	A47712	myelin/oligodendro	618	32	39.5	370	2	G30248	conserved hypotnet
546	32	39.5	218	2	A64201	DNA primase (dnase)	619	32	39.5	373	2	JC4737	G protein-coupled
547	32	39.5	220	2	H87439	hypothetical protei	620	32	39.5	374	2	JC4162	p2y receptor - bov
548	32	39.5	221	2	E83868	hypothetical prote	621	32	39.5	381	1	B69030	probable muconate
549	32	39.5	223	2	H82482	DNA-3-methyladenin	622	32	39.5	381	2	D64525	conserved hypotnet
550	32	39.5	225	2	C85475	transcription regu	623	32	39.5	382	2	F96009	hypothetical membr
551	32	39.5	230	2	C97104	hypothetical prote	624	32	39.5	383	2	AH01139	galactokinase (EC
552	32	39.5	232	2	T15146	hypothetical prote	625	32	39.5	383	2	T04641	homeobox protein 1
553	32	39.5	233	2	S75267	hypothetical prote	626	32	39.5	390	2	A70656	hypothetical prote
554	32	39.5	235	2	A10365	probable two-compo	627	32	39.5	392	1	C69851	macrolide glycosyl
555	32	39.5	240	1	QRECGQ	glutamine transpor	628	32	39.5	393	2	AB2752	enolase [imported]
556	32	39.5	240	2	A10600	glutamine transpor	629	32	39.5	394	2	C96728	hypothetical prote
557	32	39.5	240	2	G90739	hypothetical prote	630	32	39.5	397	2	H84277	orc / cell divisio
558	32	39.5	240	2	A85590	hypothetical prote	631	32	39.5	399	2	T06251	GRP binding protei
559	32	39.5	241	2	A87712	endonuclease III I	632	32	39.5	400	2	G01977	d3 dopamine recept
560	32	39.5	241	2	F81997	leucyl/phenylalany	633	32	39.5	404	2	AD2994	glycosyltransferas
561	32	39.5	247	2	A55717	myelin/oligodendro	634	32	39.5	406	2	T24301	hypothetical prote
562	32	39.5	248	2	AH3474	DNA- (apurinic or a	635	32	39.5	408	2	E75452	probable transposa
563	32	39.5	248	2	T33281	hypothetical prote	636	32	39.5	408	2	G75376	hypothetical prote
564	32	39.5	250	2	AE0621	probable prophage	637	32	39.5	410	2	S74705	hypothetical prote
565	32	39.5	252	2	JC5237	osmotin-like prote	638	32	39.5	412	2	H72741	probable thiazole
566	32	39.5	257	2	AD1996	3-methyl-2-oxobuta	639	32	39.5	413	2	C81278	hypothetical prote
567	32	39.5	260	2	D75404	phosphatidylglycer	640	32	39.5	414	2	B98354	hepsin (EC 3.4.21.
568	32	39.5	262	2	AF2698	NAD/NADP dependent	641	32	39.5	416	1	S33777	enolase (2-phospho
569	32	39.5	265	2	AH3044	conserved hypotnet	642	32	39.5	424	2	A97533	hypothetical zinc-
570	32	39.5	265	2	D98241	hypothetical prote	643	32	39.5	425	2	AC2756	MFS permease [limpo
571	32	39.5	265	2	S74449	conserved hypotnet	644	32	39.5	425	2	B84619	probable serine ca
572	32	39.5	268	2	H87474	naphthoate synthas	645	32	39.5	433	2	E84618	probable serine ca
573	32	39.5	273	2	D98873	extragenic suppress	646	32	39.5	435	2	B95338	hypothetical prote
574	32	39.5	275	2	E82554	LPS biosynthesis p	647	32	39.5	437	2	D84619	hypothetical prote
575	32	39.5	278	2	F86651	hypothetical prote	648	32	39.5	439	2	D71115	hypothetical prote
576	32	39.5	278	2	H87663	conserved hypotnet	649	32	39.5	442	2	T15310	dihydrolipoamide d
577	32	39.5	280	2	H95901	hypothetical prote	650	32	39.5	449	2	E72383	protein kinase (EC
578	32	39.5	280	2	H95901	hypothetical prote	651	32	39.5	449	2	T16891	protein kinase (EC
579	32	39.5	290	2	E84797	probable dehydroge	652	32	39.5	454	2	A38643	mRNA guanylyltrans
580	32	39.5	293	2	F97480	NADH2 dehydrogenas	653	32	39.5	455	2	S67775	glucosyl transfera
581	32	39.5	293	2	I39913	hypothetical prote	654	32	39.5	459	2	S59731	dihydrolipoamide d
582	32	39.5	300	2	T20905	acyl carrier prote	655	32	39.5	467	2	T02238	degr4 protease lik
583	32	39.5	303	2	H69551	protein kinase Sry	656	32	39.5	468	2	AG2899	Deg4 protease lik
584	32	39.5	306	2	I49068	hypothetical prote	657	32	39.5	468	2	H97674	G-protein signalin
585	32	39.5	307	2	AF2402	probable transcrip	658	32	39.5	480	2	T46925	hypothetical prote
586	32	39.5	309	2	E83243	flavoprotein oxid	659	32	39.5	481	2	S62427	protein kinase Sry
587	32	39.5	312	2	AC3077	hypothetical prote	660	32	39.5	483	2	A39676	protein kinase cik
588	32	39.5	312	2	F98209	hypothetical prote	661	32	39.5	484	2	S53641	conserved hypotnet
589	32	39.5	317	2	S63358	hypothetical prote	662	32	39.5	485	2	T39508	2-hydroxymuconic s
590	32	39.5	319	2	A12125	hypothetical prote	663	32	39.5	486	2	S10772	hydrogenase (EC 1.
591	32	39.5	321	2	T01477	protein kinase hom	664	32	39.5	486	2	H75060	gene 14 protein -
592	32	39.5	321	2	AE1282	conserved hypotnet	665	32	39.5	486	2	S10772	NADH2 dehydrogenas
593	32	39.5	321	2	A11653	conserved hypotnet	666	32	39.5	488	2	S30959	surface glycoprote
594	32	39.5	322	2	B97457	BH305 conserved h	667	32	39.5	488	2	T06260	hypothetical prote
595	32	39.5	323	2	AD2675	conserved hypotnet	668	32	39.5	489	1	VGXPMV	hypothetical prote
596	32	39.5	323	2	D86216	protein T23G18.8 I	669	32	39.5	491	2	T52398	probable transemb
597	32	39.5	324	2	B84064	glycerate dehydrog	670	32	39.5	493	2	T33569	chitinase (EC 3.2.
598	32	39.5	326	2	G98231	exoO protein (impo	671	32	39.5	504	2	A38221	NADH2 dehydrogenas
599	32	39.5	326	2	AF1054	succinoglycan bios	672	32	39.5	504	2	S28743	probable cell divi
600	32	39.5	326	2	T10166	restriction endonu	673	32	39.5	522	2	S28743	hypothetical prote
601	32	39.5	326	2	S74844	lmbp protein - Syn	674	32	39.5	524	2	F70579	CTP synthetase [im
602	32	39.5	328	2	AE0301	conserved hypotnet	675	32	39.5	531	2	T15329	hypothetical prote
603	32	39.5	332	2	A86892	glycosyltransferas	676	32	39.5	535	2	C86686	seizure-related me
604	32	39.5	332	2	A82558	dopamine receptor	677	32	39.5	542	2	T28719	probable fatty-aci
605	32	39.5	342	2	A96926	endoglucanase, ami	678	32	39.5	546	2	JC4798	related to RNA-bin
606	32	39.5	346	1	Z1BRF1	gene I protein - p	679	32	39.5	562	2	H69545	hypothetical prote
607	32	39.5	348	1	Z1BRFD	gene I protein - p	680	32	39.5	564	2	T49322	hypothetical prote
608	32	39.5	348	1	Z1BRFD	gene I protein - p	681	32	39.5	567	2	AG2008	hypothetical prote
609	32	39.5	348	1	Z1BRFD	lrm protein - Caen	682	32	39.5	570	2	T46911	protein Y54E2A.1 I
610	32	39.5	348	2	T16076	hypothetical prote	683	32	39.5	580	2	H87963	sulfite reductase
611	32	39.5	348	2	G86270	hypothetical prote	684	32	39.5	589	2	E97376	hypothetical prote
612	32	39.5	353	2	AB2483	phosphate binding	685	32	39.5	589	2	AD2594	
613	32	39.5	354	2	AC0389		686	32	39.5				

687 32 39.5 594 1 WZBER2 protein-serine/thr
688 32 39.5 596 2 A85438 hypothetical prote
689 32 39.5 597 2 A50721 hydrogenase-1 larg
690 32 39.5 598 2 I51368 gamma-aminobutyric
691 32 39.5 599 1 A46037 gamma-aminobutyric
692 32 39.5 599 1 ACR7GT gamma-aminobutyric
693 32 39.5 599 2 S11073 gamma-aminobutyric
694 32 39.5 601 2 D83583 probable acyl-CoA
695 32 39.5 602 2 H86468 protein F12K21.20
696 32 39.5 617 2 D90487 maltose ABC transp
697 32 39.5 618 2 S09251 hydrogenase (EC 1.
698 32 39.5 619 1 JH0776 hydrogenase (EC 1.
699 32 39.5 620 1 S33253 protein-tyrosine k
700 32 39.5 621 2 I38467 low density lipopr
701 32 39.5 625 1 A43030 protein-tyrosine k
702 32 39.5 640 2 S49932 MET30 protein - ye
703 32 39.5 642 2 G69786 ABC transporter (A
704 32 39.5 643 2 T19135 cholecystokinin ty
705 32 39.5 661 2 P96665 protein P22C12.14
706 32 39.5 668 2 T15305 hypothetical prote
707 32 39.5 675 2 T04644 hypothetical prote
708 32 39.5 699 2 T14904 NADPH-ferrihemopro
709 32 39.5 700 2 A84243 phosphoribosylform
710 32 39.5 700 2 A96690 hypothetical prote
711 32 39.5 703 2 AG0242 probable membrane
712 32 39.5 705 2 F70475 VacB protein (ribo
713 32 39.5 706 2 D82452 anaerobic ribonuc
714 32 39.5 713 2 T21201 hypothetical prote
715 32 39.5 719 2 A30047 enhancer of split
716 32 39.5 721 2 C71014 hypothetical prote
717 32 39.5 733 2 A58891 probable N-methylh
718 32 39.5 754 2 S50601 hypothetical prote
719 32 39.5 759 2 T41295 protein transport
720 32 39.5 763 2 E36693 probable terpene s
721 32 39.5 771 2 T01315 hypothetical prote
722 32 39.5 772 2 S35581 ESG protein - mous
723 32 39.5 784 2 T18452 hypothetical prote
724 32 39.5 788 1 B4AG58 virB4 protein prec
725 32 39.5 789 1 B4AG66 virB4 protein prec
726 32 39.5 792 2 AH3248 component of type
727 32 39.5 796 2 AG1849 serine/threonine k
728 32 39.5 797 2 B84642 hypothetical prote
729 32 39.5 829 2 T32744 hypothetical prote
730 32 39.5 835 2 T26086 hypothetical prote
731 32 39.5 848 2 B89042 protein F14F9.3 [i
732 32 39.5 857 2 B84182 hypothetical prote
733 32 39.5 882 2 S41034 taste receptor T1R
734 32 39.5 882 2 S41034 hypothetical prote
735 32 39.5 888 2 S64016 probable regulator
736 32 39.5 901 2 G1286 probable pyruvate,
737 32 39.5 951 2 T08987 probable cadmium-t
738 32 39.5 965 2 S62935 hypothetical prote
739 32 39.5 977 2 H84469 hypothetical prote
740 32 39.5 985 1 DJBE11 DNA-directed DNA p
741 32 39.5 1034 2 JC5569 serine proteinase
742 32 39.5 1050 2 JW0092 serine-threonine k
743 32 39.5 1052 2 T04439 hypothetical prote
744 32 39.5 1086 2 T24214 hypothetical prote
745 32 39.5 1086 2 T40354 hypothetical prote
746 32 39.5 1127 2 G71274 hypothetical prote
747 32 39.5 1137 2 T18625 atrial natriuretic
748 32 39.5 1147 2 T35781 hypothetical prote
749 32 39.5 1189 2 T51491 hypothetical prote
750 32 39.5 1191 2 T13850 gene u-shaped prot
751 32 39.5 1246 2 T51085 related to protein
752 32 39.5 1251 2 JH0256 botulinum neurotox
753 32 39.5 1252 2 S21178 botulinum neurotox
754 32 39.5 1314 2 G02870 KRAA0197 protein -
755 32 39.5 1318 2 T39066 hypothetical prote
756 32 39.5 1326 2 H89134 protein F25G6.9 [i
757 32 39.5 1407 2 S59823 probable membrane
758 32 39.5 1420 2 T02644 ABC-type transport
759 32 39.5 1449 2 B84426 hypothetical prote

760 32 39.5 1459 2 T24088 protein-tyrosine k
761 32 39.5 1520 1 TVFPA hypothetical prote
762 32 39.5 1545 2 P98262 glutamate synthase
763 32 39.5 1581 2 A83022 two-component hybr
764 32 39.5 1627 2 AE2109 protein C28H8.3 [i
765 32 39.5 1722 2 A88470 tag24 protein - mo
766 32 39.5 1944 2 A55117 sodium channel pro
767 32 39.5 1957 2 S68453 genome polyprotein
768 32 39.5 2194 1 GNNYB7 notch3 protein - h
769 32 39.5 2318 2 S45306 notch3 protein - h
770 32 39.5 2321 2 S78549 hypothetical prote
771 32 39.5 3507 2 T34513 trithorax protein
772 32 39.5 3759 2 A35085 perlecan precursor
773 32 39.5 4391 2 A38096 hypothetical prote
774 31.5 38.9 113 2 T44561 probable carbonyl
775 31.5 38.9 173 2 AD3517 hypothetical prote
776 31.5 38.9 236 2 T33589 GTP-binding regula
777 31.5 38.9 290 2 T02300 GTP-binding protei
778 31.5 38.9 325 2 T06784 probable GTP-bind
779 31.5 38.9 325 2 T09613 GTP-binding protei
780 31.5 38.9 328 2 T16370 fructose-bisphosph
781 31.5 38.9 349 2 C72097 1,6-fructose biph
782 31.5 38.9 349 2 A86526 hypothetical prote
783 31.5 38.9 450 2 D86356 hypothetical prote
784 31.5 38.9 670 2 G96790 transferrin precur
785 31.5 38.9 695 2 S49163 VirB4 type IV secr
786 31.5 38.9 792 2 D95351 glucose dehydrogen
787 31.5 38.9 801 1 S00943 conserved hypothet
788 31.5 38.9 881 2 F69438 probable cellulose
789 31.5 38.9 1088 2 H84604 hypothetical thol
790 31.5 38.9 1275 2 T41523 hypothetical prote
791 31.5 38.9 1599 2 T16210 hypothetical prote
792 31 38.3 18 2 S33645 hypothetical prote
793 31 38.3 61 2 E90761 hypothetical prote
794 31 38.3 63 2 D97714 hypothetical prote
795 31 38.3 64 2 T21841 hypothetical prote
796 31 38.3 69 1 D64989 heme exporter prot
797 31 38.3 69 2 G91014 heme exporter prot
798 31 38.3 69 2 A85859 heme exporter prot
799 31 38.3 80 2 AC2394 hypothetical prote
800 31 38.3 85 2 A85625 ribosomal protein
801 31 38.3 89 2 G64381 hypothetical prote
802 31 38.3 102 2 S49794 hypothetical prote
803 31 38.3 108 2 T16555 hypothetical prote
804 31 38.3 112 2 C71221 phospholipase A2 h
805 31 38.3 121 1 FC4024 hypothetical prote
806 31 38.3 123 2 D96496 hypothetical prote
807 31 38.3 128 2 AC3164 hypothetical prote
808 31 38.3 131 2 T28801 hypothetical prote
809 31 38.3 132 2 I40566 histone H2A.1 - Te
810 31 38.3 137 1 HST591 hypothetical prote
811 31 38.3 137 2 B72786 histone H2A.1 - Te
812 31 38.3 138 2 S41471 hypothetical prote
813 31 38.3 140 2 T36484 hypothetical prote
814 31 38.3 140 2 F81446 hemoglobin theta-1
815 31 38.3 141 1 HTOR ribosomal protein
816 31 38.3 141 2 H71975 ribosomal protein
817 31 38.3 141 2 D64530 Spec2d protein - s
818 31 38.3 141 2 S01772 hemoglobin alpha c
819 31 38.3 142 1 HANE hemoglobin theta-1
820 31 38.3 142 1 HTHU hemoglobin theta-1
821 31 38.3 142 2 A27792 hypothetical prote
822 31 38.3 146 2 B82700 hypothetical prote
823 31 38.3 159 2 B90361 hypothetical prote
824 31 38.3 160 2 S76328 flm3-region hypoth
825 31 38.3 160 2 A53293 glycine cleavage s
826 31 38.3 163 2 T12561 hypothetical prote
827 31 38.3 163 2 D39925 conserved hypothet
828 31 38.3 171 2 AH2855 hypothetical prote
829 31 38.3 171 2 F97632 HD family hydrolas
830 31 38.3 173 2 F97105 transcription regu
831 31 38.3 174 2 C82396 oxido-reductase wit
832 31 38.3 177 2 AB3221

833	31	38.3	178	2	PT0219	Ig kappa chain V-C	906	31	38.3	292	2	G95926	probable saccharid
834	31	38.3	180	2	S06616	chorion protein s1	907	31	38.3	293	2	AE3155	oxidoreductase Atu
835	31	38.3	181	2	H72855	fibroblast growth	908	31	38.3	295	2	C83244	conserved hypothet
836	31	38.3	183	2	H72572	hypothetical prote	909	31	38.3	295	2	A44064	hepatic leukemia f
837	31	38.3	183	2	AD2549	hypothetical prote	910	31	38.3	295	2	S58525	L-serine dehydrata
838	31	38.3	188	2	G95889	probable oxidoredu	911	31	38.3	295	2	H83961	hypothetical prote
839	31	38.3	188	2	AD0945	probable DNA-bindi	912	31	38.3	298	2	D70880	hypothetical prote
840	31	38.3	192	2	C71602	hypothetical prote	913	31	38.3	299	1	D70181	protein-export mem
841	31	38.3	196	2	I38022	hypothetical prote	914	31	38.3	300	2	F75586	urea/short-chain a
842	31	38.3	198	2	C32998	chorion protein S1	915	31	38.3	300	2	C70586	probable bex - Myc
843	31	38.3	199	2	C83761	hypothetical prote	916	31	38.3	309	2	C47057	homoserine kinase
844	31	38.3	203	2	T36403	probable tetR-fami	917	31	38.3	312	1	S46047	probable 3-methyl-
845	31	38.3	205	2	AH1907	transcription regu	918	31	38.3	313	2	D98152	hypothetical prote
846	31	38.3	206	2	F72670	hypothetical prote	919	31	38.3	324	2	AB0982	probable 2-hydroxy
847	31	38.3	208	2	G69783	hypothetical prote	920	31	38.3	326	2	S59101	NADH2 dehydrogenas
848	31	38.3	208	2	AH2535	hypothetical prote	921	31	38.3	328	2	B86030	probable dehydroge
849	31	38.3	209	2	S32681	DNA-directed RNA p	922	31	38.3	328	2	C65154	probable 2-hydroxy
850	31	38.3	210	2	E97843	DNA-(apurinic or a	923	31	38.3	328	2	F91183	probable dehydroge
851	31	38.3	211	2	A05123	monofunctional bio	924	31	38.3	331	1	A54932	zeta-crystallin /
852	31	38.3	211	2	AB3286	hypothetical prote	925	31	38.3	331	2	AG3140	hypothetical prote
853	31	38.3	211	2	T34501	endonuclease III (926	31	38.3	331	2	E98147	NiFR3/Smml family
854	31	38.3	212	2	F71634	phosphoglycerate m	927	31	38.3	331	2	B82341	hypothetical prote
855	31	38.3	212	2	AH1944	hypothetical prote	928	31	38.3	333	2	T33311	hypothetical prote
856	31	38.3	214	2	I69804	hypothetical prote	929	31	38.3	334	2	T41713	hypothetical prote
857	31	38.3	214	2	B86639	acyl carrier prote	930	31	38.3	335	2	T25498	hypothetical prote
858	31	38.3	215	2	G82564	3-isopropylmalate	931	31	38.3	336	2	AE2525	hypothetical prote
859	31	38.3	218	2	E85021	hypothetical prote	932	31	38.3	339	2	E83496	GTP-binding regula
860	31	38.3	218	2	H97059	hypothetical prote	933	31	38.3	340	1	RGFFBH	hypothetical prote
861	31	38.3	220	2	A95956	hypothetical prote	934	31	38.3	341	2	C93067	agrocipoline synth
862	31	38.3	221	2	B69861	alkaline phosphata	935	31	38.3	344	2	AE3228	hypothetical prote
863	31	38.3	223	2	G82427	hypothetical prote	936	31	38.3	345	2	S51948	hypothetical prote
864	31	38.3	224	2	AH1815	two-component resp	937	31	38.3	345	1	B46113	protein kinase (EC
865	31	38.3	227	2	T12797	immunity protein d	938	31	38.3	349	1	TVCHMS	spore coat protein
866	31	38.3	229	2	D90668	probable xanthine	939	31	38.3	351	2	I39808	conserved hypothet
867	31	38.3	229	2	F64754	probable oxidoredu	940	31	38.3	357	2	A72280	probable aminotran
868	31	38.3	229	2	G85518	glucuronolactone r	941	31	38.3	359	2	E83251	lipid A disacchari
869	31	38.3	229	2	T40439	hypothetical prote	942	31	38.3	360	2	C84628	lipid-a-disacchari
870	31	38.3	231	2	G83565	probable bacteriop	943	31	38.3	360	2	D71888	hypothetical prote
871	31	38.3	232	2	S28703	precorrin-2 methyl	944	31	38.3	360	2	AE1931	3-dehydroquinat s
872	31	38.3	232	2	B69046	hypothetical prote	945	31	38.3	362	2	AE0019	protein C34G6.5 [i
873	31	38.3	232	2	A87612	petR protein, prob	946	31	38.3	362	2	B87789	hypothetical prote
874	31	38.3	233	2	B64436	cobalamin biosynth	947	31	38.3	363	2	T22255	hypothetical prote
875	31	38.3	233	2	E86223	hypothetical prote	948	31	38.3	364	2	T23819	uroporphyrinogen d
876	31	38.3	237	2	T25955	hypothetical prote	949	31	38.3	367	2	A24411	uroporphyrinogen d
877	31	38.3	238	2	S76860	amino-acid ABC tra	950	31	38.3	367	2	G02786	uroporphyrinogen d
878	31	38.3	241	2	C71725	probable nucleosid	951	31	38.3	367	2	T10088	hydroxyproline-ric
879	31	38.3	245	2	F83270	trypsin (EC 3.4.21	952	31	38.3	368	2	C29356	hypothetical prote
880	31	38.3	247	1	B25852	carboxylesterase (953	31	38.3	369	2	T24022	nonstructure prote
881	31	38.3	247	1	JC1374	hypothetical prote	954	31	38.3	372	2	JC1069	hypothetical prote
882	31	38.3	247	2	B83816	glucose 1-dehydrog	955	31	38.3	372	2	C96631	alanine dehydrogen
883	31	38.3	248	2	F69868	hypothetical prote	956	31	38.3	372	2	T25629	hypothetical prote
884	31	38.3	248	2	H83111	tRNA pseudouridine	957	31	38.3	374	2	C82142	hypothetical prote
885	31	38.3	248	2	A99183	conserved hypothet	958	31	38.3	377	2	T00643	zinc metalloprotei
886	31	38.3	251	2	F83101	endonuclease (EC 3	959	31	38.3	377	2	AB1552	N-acetylglucosamin
887	31	38.3	252	2	S31147	hypothetical prote	960	31	38.3	377	2	AD1194	8-amino-7-oxononan
888	31	38.3	253	2	S74480	26S proteasome reg	961	31	38.3	380	2	H81830	probable rubredoxi
889	31	38.3	257	2	S56108	tributylin esteras	962	31	38.3	380	2	C81194	hypothetical prote
890	31	38.3	259	2	E95071	probable transcrip	963	31	38.3	384	2	G82976	probable phosphose
891	31	38.3	259	2	G95890	tributylin esteras	964	31	38.3	384	2	S74806	hypothetical prote
892	31	38.3	259	2	C97939	probable methyltra	965	31	38.3	389	2	S67450	protein kinase (EC
893	31	38.3	261	2	AD0171	nadh oxidoreductas	966	31	38.3	393	1	TVB666	hypothetical prote
894	31	38.3	261	2	D71838	calcium sensor pro	967	31	38.3	394	2	T31891	cytochrome b-like
895	31	38.3	268	2	I53413	hypothetical prote	968	31	38.3	396	2	S49592	probable YBR162c ho
896	31	38.3	268	2	H82757	probable haloacid	969	31	38.3	396	2	S56954	hypothetical prote
897	31	38.3	269	2	AC0497	probable arylmalon	970	31	38.3	398	2	C84579	glycine hydroxymet
898	31	38.3	272	2	H72484	gag protein - frui	971	31	38.3	399	2	S34681	hypothetical prote
899	31	38.3	273	2	S00953	conserved hypothet	972	31	38.3	400	2	S30382	streptomycin biosy
900	31	38.3	275	2	C87634	homeotic protein H	973	31	38.3	402	2	B95354	hypothetical prote
901	31	38.3	284	2	A75356	hypothetical prote	974	31	38.3	403	2	C75405	serine hydroxymeth
902	31	38.3	285	2	A43556	probable LysR-fami	975	31	38.3	406	2	G84774	serine hydroxymeth
903	31	38.3	285	2	T15252		976	31	38.3	412	2	F90004	
904	31	38.3	291	2	A95392		977	31	38.3	413	2	E84120	
905	31	38.3					978	31	38.3				

979 31 38.3 413 2 AC1392 glycine hydroxymet
980 31 38.3 413 2 AB1767 glycine hydroxymet
981 31 38.3 414 1 J01016 glycine hydroxymet
982 31 38.3 414 2 HA1383 glycine hydroxymet
983 31 38.3 414 2 S18530 hypohetical prote
984 31 38.3 415 1 I40483 glycine hydroxymet
985 31 38.3 417 1 S00845 hepsin (EC 3.4.21.
986 31 38.3 417 1 G84963 glycine hydroxymet
987 31 38.3 417 2 H70174 glycine hydroxymet
988 31 38.3 419 2 T15199 hypohetical prote
989 31 38.3 421 2 T43534 transcription fact
990 31 38.3 421 2 T14773 hypohetical prote
991 31 38.3 423 1 DBBYMC malate dehydrogena
992 31 38.3 425 1 KXKL2A keratin, 64K type
993 31 38.3 427 2 F72341 glycine hydroxymet
994 31 38.3 427 2 S75210 glycine hydroxymet
995 31 38.3 427 2 AF2406 serine hydroxymeth
996 31 38.3 427 2 I51580 XPX2 protein - Af
997 31 38.3 430 2 S12705 site-specific DNA-
998 31 38.3 430 2 J01050 deoxyfructosyl-glu
999 31 38.3 430 2 I39727 mannopine biosynth
1000 31 38.3 434 2 D72390 conserved hypothet

ALIGNMENTS

RESULT 1
A32378
probable ATP-dependent RNA helicase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change 09-Jul-2004
C:Accession: A32378
R:ieroy, P.; Alzari, P.; Sassoon, D.; Wolgemuth, D.; Fellous, M.
Cell 57, 549-559, 1989
A:Title: The protein encoded by a murine male germ cell-specific transcript is a putative
A:Reference number: A32378; MUID:89249320; PMID:2720782
A:Accession: A32378
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-660 <LER>
A:Cross-references: UNIPROT:P16381; UNIPARC:UPI0000003F94; GB:J04847; NID:g200388; PIDN:
C:Superfamily: ATP-dependent RNA helicase DBP1
C:Keywords: ATP; DNA binding; nucleotide binding; P-loop
F:223-230/Region: nucleotide-binding motif A (P-loop)
F:342-347/Region: nucleotide-binding motif B
F:346-349/Region: DEAD motif

Query Match 59.3%; Score 48; DB 2; Length 660;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
||:|||||
Db 455 SLEDFLYHEGYAC 467

RESULT 2
184741
RNA helicase - mouse
N:Alternate names: RNA helicase ERH
C:Species: Mus musculus (house mouse)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: I84741; S56112
R:Gee, S.L.; Conboy, J.G.
Gene 140, 171-177, 1994
A:Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting h
A:Reference number: 149731; MUID:94192995; PMID:8144024
A:Accession: I84741
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-662 <RES>
A:Cross-references: UNIPROT:Q62167; UNIPARC:UPI00000044D80; GB:L25126; NID:g407995; PIDN:

A:Genetics: RES1
R:Snowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.
Biochem. J. 308, 839-846, 1995
A:Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of R
A:Reference number: S56112; MUID:97104282; PMID:8948440
A:Accession: S56112
A:Molecule type: mRNA
A:Residues: 1-229 'R', 231-662 <SOW>
A:Cross-references: UNIPARC:UPI0000172EBF; GB:Z38117; NID:g1835121; PID:g1835122
A:Experimental source: strain c57bl/6; notochord
A:Genetics: SOW1
A:Note: the sequence is revised in GenBank entry MMDBRNAHL, release 117, (PID:1835122)
C:Genetics: <RES1>
A:Gene: Eif4a-rsl; MGI:Ddx19
A:Cross-references: MGI:995526
C:Genetics: <SOW1>
A:Gene: MGI:Ddx3
A:Cross-references: MGI:103064
A:Map position: 1
C:Superfamily: ATP-dependent RNA helicase DBP1
C:Keywords: ATP; nucleotide binding; P-loop
F:224-231/Region: nucleotide-binding motif A (P-loop)
F:343-348/Region: nucleotide-binding motif B
F:347-350/Region: DEAD motif

Query Match 59.3%; Score 48; DB 1; Length 662;
Best Local Similarity 61.5%; Pred. No. 3.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
||:|||||
Db 456 SLEDFLYHEGYAC 468

RESULT 3
A81710
ribosomal protein S1 TC0373 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPN
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81710
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOPN and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81710
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <TET>
A:Cross-references: UNIPROT:P38016; UNIPARC:UPI00000578D3; GB:AE002304; GB:AE002160; NID:
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0373
C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 55.6%; Score 45; DB 2; Length 536;
Best Local Similarity 63.6%; Pred. No. 8.8;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGY 12
||:|||||
Db 512 SIKFLAHGGH 522

RESULT 4
S13654
ATP-dependent RNA helicase - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S13654; S29676
R:Gururajan, R.; Perry-O'Keefe, H.; Melton, D.A.; Weeks, D.L.
Nature 349, 717-719, 1991

A;Title: The Xenopus localized messenger RNA An3 may encode an ATP-dependent RNA helicase
A;Reference number: S13654; MUID:91141586; PMID:1996140
A;Accession: S13654
A;Molecule type: mRNA
A;Residues: 1-697 <GUR>
A;Cross-references: UNIPROT:P24346; UNIPARC:UPI0000125AC2; EMBL:X57328; NID:965059; PIDN:19954
C;Superfamily: ATP-dependent RNA helicase DBP1
C;Keywords: ATP; nucleotide binding; P-loop
F;260-272/Region: ATP binding #status predicted
F;285-272/Region: nucleotide-binding motif A (P-loop)
F;384-389/Region: nucleotide-binding motif B
F;386-398/Region: ATP binding #status predicted
F;388-391/Region: DEAD motif

Query Match 55.6%; Score 45; DB 1; Length 697;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

QY 2 SLOEFLSHGGYVC 14
Db 497 ALEDFLYHEGYAC 509
:::|||||

RESULT 5
JC5401
inositol 1,3,4-trisphosphate 5/6-kinase (EC 2.7.-.-) [imported] - Arabidopsis thaliana
N;Alternate names: protein F5E19 100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C;Accession: JC5401; T51508; T51954
R;Wilson, M.P.; Majerus, P.W.
Biochem. Biophys. Res. Commun. 232, 678-681, 1997
A;Title: Characterization of a cDNA encoding Arabidopsis thaliana inositol 1,3,4-trisphosphate 5/6-kinase
A;Reference number: JC5401; MUID:97271384; PMID:9126335
A;Accession: JC5401
A;Molecule type: mRNA
A;Residues: 1-319 <WIL>
A;Cross-references: UNIPROT:O81633; UNIPARC:UPI000009EB2A; GB:AF080173; NID:93396078; PMID:93396078; PIDN:93396078
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: 225394
A;Accession: T51508
A;Molecule type: DNA
A;Residues: 1-196, 'K', 198-319 <SAT>
A;Cross-references: UNIPARC:UPI00000ACC41; EMBL:AL391147; NID:99755718; PIDN:CAC01840.1;
A;Experimental source: cultivar Columbia; BAC clone F5E19
C;Genetics:
A;Map position: 5
A;Note: F5E19 100
C;Keywords: phosphotransferase

Query Match 54.3%; Score 44; DB 2; Length 319;
Best Local Similarity 53.6%; Pred. No. 7.7;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LOEFLSHGGYV 13
Db 176 LOEFVNHGGVI 186
|||||

RESULT 6
A36220
transforming protein (s-myc) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 09-Jul-2004
C;Accession: A36220
R;Sugiyama, A.; Kume, A.; Nemoto, K.; Lee, S.Y.; Asami, Y.; Nemoto, F.; Nishimura, S.; K
Proc. Natl. Acad. Sci. U.S.A. 86, 9144-9148, 1989
A;Title: Isolation and characterization of 8-myc, a member of the rat myc gene family.
A;Reference number: A36220; MUID:90083224; PMID:2594755
A;Accession: A36220
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-429 <SUG>
A;Cross-references: UNIPROT:P23999; UNIPARC:UPI000012FAFE; GB:M29069; NID:9205553; PIDN:19954
C;Superfamily: myc transforming protein; myc transforming protein homology
C;Keywords: DNA binding; leucine zipper; nucleus; phosphoprotein
F;19-429/Domain: myc transforming protein homology <MYC>
F;398-426/Region: leucine zipper motif

Query Match 54.3%; Score 44; DB 2; Length 429;
Best Local Similarity 56.7%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SLOEFLSHGGYV 13
Db 227 SLEDFLSNGYV 238
|||||

RESULT 7
T19954
hypothetical protein C45B11.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004
C;Accession: T19954
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19202
A;Accession: T19954
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-293 <WIL>
A;Cross-references: UNIPROT:Q18639; UNIPARC:UPI0000077952; EMBL:274029; PIDN:CAA98431.1;
A;Experimental source: clone C45B11
C;Genetics:
A;Gene: CESP:C45B11.3
A;Map position: 5
A;Introns: 9/2; 66/3
C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 51.9%; Score 42; DB 2; Length 293;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLOEFLSHGG 11
Db 208 SAMEFLSHGG 217
|||||

RESULT 8
T10544
inositol 1,3,4-trisphosphate 5/6-kinase homolog T12G13.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10544
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10544
A;Molecule type: DNA
A;Residues: 1-338 <BEV>
A;Cross-references: UNIPROT:Q9SUG3; UNIPARC:UPI0000048910; EMBL:AL080252; GSPDB:GN00062,
A;Experimental source: cultivar Columbia; BAC clone T12G13
C;Genetics:
A;Gene: ATSP:T12G13.10
A;Map position: 4
A;Introns: 65/2; 92/3; 103/3; 144/1; 175/1; 215/1; 274/1; 292/3

Query Match 51.9%; Score 42; DB 2; Length 338;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LOEFLSHGG 11
Db 208 LOEFVNHGG 216
|||||

RESULT 9

glutamine synthetase - Thermotoga maritima (strain MSB8)
C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: B72313
R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing

A/Reference number: A72200; MUID:99287316; PMID:10360571

A/Accession: B72313

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-439 <ARN>

A/Cross-references: UNIPROT:P36205; UNIPARC:UPI000012B707; GB:AE001758; GB:AE000512; NID

A/Experimental source: strain MSB8

C/Genetics:

A/Gene: TM0943

C/Superfamily: glutamate-ammonia ligase

Query Match 51.9%; Score 42; DB 2; Length 439;

Best Local Similarity 87.5%; Pred. No. 24;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EFLSHGGY 12

|||||

Db 145 EFLDHGGY 152

|||||

RESULT 10

H90876

probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C/Accession: H90876

R/Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: H90876

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-113 <HAY>

A/Cross-references: UNIPROT:Q8X210; UNIPARC:UPI00000D2AC7; GB:BA0000007; PIDN:BA035407.1;

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A/Gene: ECe1984

C/Superfamily: phage lambda minor tail protein M

Query Match 51.2%; Score 41.5; DB 2; Length 113;

Best Local Similarity 81.8%; Pred. No. 7.3;

Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 3 LQEFLS-HGGY 12

|||||

Db 62 LEEFLSMHGGY 72

|||||

RESULT 11

T23345

hypothetical protein K05C4.11 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T23345

R/Harris, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19729

A/Accession: T23345

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-436 <WIL>

A/Cross-references: UNIPROT:Q9XU02; UNIPARC:UPI0000076837; EMBL:Z81564; PIDN:CAB04576.1;

A/Experimental source: clone K05C4

C/Genetics:

A/Gene: CRSP:K05C4.11

A/Map position: 1

A/Introns: 34/1; 146/1; 174/1; 333/3; 369/1

C/Superfamily: Caenorhabditis elegans protein K05C4.11; LDL receptor ligand-binding repeat

F;330-369/Domain: LDL receptor ligand-binding repeat homology <LDL>

Query Match 50.6%; Score 41; DB 2; Length 436;

Best Local Similarity 60.0%; Pred. No. 35;

Matches 6; Conservative 2; Mismatches 0; Gaps 0;

Qy 5 EFLSHGGYVC 14

|||||

Db 321 QFMCHGGQVC 330

|||||

RESULT 12

I38026

MLN 62 protein - human

C/Species: Homo sapiens (man)

C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 07-Jul-2003

C/Accession: I38026; S60681

R/Tomasetto, C.; Regnier, C.H.; Moog-Lutz, C.; Mattei, M.G.; Chenard, M.P.; Lidereau, R.

Genomics 28, 367-376, 1995

A/Title: Identification of four novel human genes amplified and overexpressed in breast

A/Reference number: I37080; MUID:96039245; PMID:7490069

A/Accession: I38026

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-470 <RES>

A/Cross-references: UNIPARC:UPI0000073E24; EMBL:X80200; NID:9951276; PIDN:CAA56491.1-PI

A/Note: submitted to the EMBL Data Library, July 1994

C/Genetics:

A/Gene: MLN62; CART1

C/Superfamily: TNF receptor-associated factor (TRAF); RING finger homology

F;14-63/Domain: RING finger homology <RRN>

Query Match 50.6%; Score 41; DB 2; Length 470;

Best Local Similarity 66.7%; Pred. No. 38;

Matches 8; Conservative 1; Mismatches 0; Gaps 0;

Qy 3 LQEFLSHGGYVC 14

|||||

Db 43 LQEFLSRGVFKC 54

|||||

RESULT 13

G88925

protein T22H9.2 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C/Accession: G88925

R/anonymous, the C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: G88925

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-880 <STO>

A/Cross-references: UNIPROT:Q9TXN6; UNIPARC:UPI0000080264; GB:chr_V; PIDN:AAC69227.1; PI

C/Genetics:

A/Gene: T22H9.2

A/Map position: 5

Query Match 50.0%; Score 40.5; DB 2; Length 880;

Best Local Similarity 40.0%; Pred. No. 87;

Matches 8; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Qy 2 SLQEFLSH-----GGYVC 14
:|||||:
135 NLDEFFTHYHQNGGYLC 154

RESULT 14
PQ0446
DNA-binding protein B2F - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PQ0446; JQ1679; S27764
R:Perisic, O.; Lam, E.
Plant Cell 4, 831-838, 1992
A:Title: A tobacco DNA binding protein that interacts with a light-responsive box II element
A:Reference number: PQ0446; MUID:93005727; PMID:1392597
A:Accession: PQ0446
A:Molecule type: mRNA
A:Residues: 1-392 <PER>
A:Cross-references: UNIPROT:Q43326; UNIPARC:UPI00000ACFP45; GB:S46942; NID:g258167; PIDN:
R:Experimental source: leaf, strain SR1
R:Gilmarin, P.M.; Memelink, J.; Hiratsuka, K.; Kay, S.A.; Chua, N.H.
Plant Cell 4, 839-849, 1992
A:Title: Characterization of a gene encoding a DNA binding protein with specificity for
A:Reference number: JQ1679; MUID:93005728; PMID:1392598
A:Accession: JQ1679
A:Molecule type: mRNA
A:Residues: 13-392 <GIL>
A:Cross-references: UNIPARC:UPI000016DE7F; GB:M93436; NID:gi70270; PIDN:AAA34085.1; PID:
C:Keywords: DNA binding
F:75-103/Domain: alpha-helical rod <RD1>
F:115-124/Domain: alpha-helical rod <RD2>
F:138-150/Domain: alpha-helical rod <RD3>
F:159-174/Domain: alpha-helical rod <RD4>

Query Match 49.4%; Score 40; DB 2; Length 392;
Best Local Similarity 54.5%; Pred. No. 47; Mismatches 3; Indels 2; Gaps 0;
Matches 6; Conservative 3;

Qy 4 QEFLSHGYYVC 14
:|||||:
356 RDLFSLRGTWC 366

RESULT 15
A46394
suppressor protein SSL1 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein L1531; protein YLR005W
C:Species: Saccharomyces cerevisiae
C:Date: 18-May-1994 #sequence revision 19-Jul-1996 #text_change 31-Dec-2004
C:Accession: A46394; S64827
R:Yoon, H.; Miller, S.P.; Pabich, E.K.; Donahue, T.F.
Genes Dev. 6, 2463-2477, 1992
A:Title: SSL1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for transl
A:Reference number: A46394; MUID:94040711; PMID:1340463
A:Accession: A46394
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <YOO>
A:Cross-references: UNIPROT:Q04673; UNIPARC:UPI0000053049; GB:Z17385; NID:g2695; PID:g26
R:Vandenbol, M.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64742
A:Accession: S64827
A:Molecule type: DNA
A:Residues: 1-461 <VAN>
A:Cross-references: UNIPARC:UPI0000053049; EMBL:Z73177; NID:gl360293; PID:gl360294; MIPS
A:Note: experimental_source strain S288C
C:Genetics:
A:Gene: SGD:SSL1
A:Cross-references: SGD:S0003995; MIPS:YLR005W
A:Map position: 12R

C:Superfamily: TFIIF basal transcription factor complex, subunit SSL1
C:Keywords: transmembrane protein
F:356-372/Domain: transmembrane #status predicted <TMM>

Query Match 49.4%; Score 40; DB 2; Length 461;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CSLQEFSLSHGGYVC 14
|||:-|||:
336 CSCHSKLVGGYFC 349

RESULT 16
A83884

L-arabinose isomerase araA [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: A83884

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: A83884

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-497 <STO>

A:Cross-references: UNIPROT:Q9KBQ2; UNIPARC:UPI0000125D50; GB:AP001513; GB:BA0000004; NIT

A:Experimental source: strain C-125

C:Genetics:

A:Gene: araA

C:Superfamily: L-arabinose isomerase

Query Match 49.4%; Score 40; DB 2; Length 497;
Best Local Similarity 63.6%; Pred. No. 59;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHGGY 12
:|||||:
262 ALKEFLFEGGY 272

RESULT 17
A71556

Probable S1 ribosomal protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004

C:Accession: A71556

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra

A:Reference number: A71570; MUID:99000809; PMID:9784136

A:Accession: A71556

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-569 <ARN>

A:Cross-references: UNIPROT:O84100; UNIPARC:UPI0000047BDA; GB:AE001284; GB:AE001273; NIT

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: rsl

C:Superfamily: Escherichia coli ribosomal protein S1

Query Match 49.4%; Score 40; DB 2; Length 569;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHGG 11
:|||||:
546 SIKEFLVHGG 555

RESULT 18
T27429

hypothetical protein Y79H2A.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27429
R;Matthews, L.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20365
A;Accession: T27429
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-643 <WIL>
A;Cross-references: UNIPROT:Q95QC4; UNIPARC:UPI000017BCBE; EMBL:AL110501; NID:e1542357;
A;Experimental source: clone Y79H2A
C;Genetics:
A;Gene: CESP:Y79H2A.11
A;Introns: 63/1; 365/2; 443/2; 502/2; 562/3; 611/3

Query Match 49.4%; Score 40; DB 2; Length 643;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
|:|:| |||||
Db 274 SVDQFEDGGYVC 286

RESULT 19
D56695
transducin-like enhancer-of-split homolog TLE-3 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 05-Oct-2004
C;Accession: D56695
R;Stifani, S.; Blaumueller, C.M.; Redhead, N.J.; Hill, R.E.; Artavanis-Tsakonas, S.
Nature Genet. 2, 119-127, 1992
A;Title: Human homologs of a Drosophila enhancer of split gene product define a novel fa
A;Reference number: A56695; MUID:93265135; PMID:1303260
A;Accession: D56695
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-772 <STI>
A;Cross-references: UNIPROT:Q04726; UNIPARC:UPI0000163B5B; GB:M99438; NID:g307513; PIDN:
C;Genetics:
A;Gene: GDB:TLE3; ESG: ESG3
A;Cross-references: GDB:228049; OMTM:600190
A;Map position: 15pter-15qter
C;Keywords: nucleus
F;482-514/Domain: WD repeat homology <WD1>
F;528-561/Domain: WD repeat homology <WD2>
F;614-647/Domain: WD repeat homology <WD3>
F;696-729/Domain: WD repeat homology <WD4>
F;734-770/Domain: WD repeat homology <WD5>

Query Match 49.4%; Score 40; DB 2; Length 772;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LSHGGYVC 14
|:|:| |||||
Db 483 LSHGGYVC 490

RESULT 20
D90854
minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D90854
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90854

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <HAY>
A;Cross-references: UNIPROT:Q8X5F5; UNIPARC:UPI000000D2966; GB:BA000007; PIDN:BAB35227.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1804
C;Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 LQEFLS-HGGY 12
|:|:| |||||
Db 62 LDEFSLWHGGY 72

RESULT 21
G90908
probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: G90908
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: G90908
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <HAY>
A;Cross-references: UNIPROT:Q8X5F5; UNIPARC:UPI000000D2966; GB:BA000007; PIDN:BAB35662.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2239
C;Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 3 LQEFLS-HGGY 12
|:|:| |||||
Db 62 LDEFSLWHGGY 72

RESULT 22
B90823
probable minor tail protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B90823
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <HAY>
A;Cross-references: UNIPROT:Q8X3G6; UNIPARC:UPI000000D298C; GB:BA000007; PIDN:BAB34977.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1554
C;Superfamily: phage lambda minor tail protein M

Query Match 48.8%; Score 39.5; DB 2; Length 113;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 3 LOEFLSHGGY 12
    |||||
Db 62 LDEFLSWHGGY 72

RESULT 23
C95858
probable oxidoreductase protein SMB20131 (imported) - Sinorhizobium meliloti (strain 102)
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 31-Dec-2004
C:Accession: C95858
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-Kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: C95858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <KUR>
A:Cross-references: UNIPROT:Q92X31; UNIPARC:UPI00000CB411; GB:AL591985; PIDN:CAC48531.1
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20131
A:Genome: plasmid
C:Superfamily: oxidoreductase iron-sulfur binding subunit, IorA type; ferredoxin [2Fe-2S]
C:Keywords: 2Fe-2S; metalloprotein
F:41,46,49,61/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 48.1%; Score 39; DB 2; Length 161;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 QBEFLSHGGY 14
    |||||
Db 91 QKMEHGGAC 101

RESULT 24
B75576
oxidoreductase, iron-sulfur subunit - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Dec-2004
C:Accession: B75576
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75576
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-187 <WHI>
A:Cross-references: UNIPROT:Q9RY85; UNIPARC:UPI00000D3BDD; GB:AE001863; PIDN:10567266
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0233
A:Map position: 2
C:Superfamily: oxidoreductase iron-sulfur binding subunit, IorA type; ferredoxin [2Fe-2S]
C:Keywords: 2Fe-2S; metalloprotein
F:56,61,64,76/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 48.1%; Score 39; DB 2; Length 187;
Best Local Similarity 54.5%; Pred. No. 33;

QY 3 LOEFLSHGGY 12
    |||||
Db 62 LDEFLSWHGGY 72

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 QBEFLSHGGY 14
    |||||
Db 105 QAFIEHDGYC 115

RESULT 25
F71804
probable transcription regulator - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: F71804; A71826
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: F71804
A:Molecule type: DNA
A:Residues: 1-213 <ARN>
A:Cross-references: UNIPROT:Q9Z371; UNIPARC:UPI00000D3829; GB:AE001567; PIDN:9923682
A:Experimental source: strain J99
A:Accession: A71826
A:Molecule type: DNA
A:Residues: 1-213 <AR2>
A:Cross-references: UNIPARC:UPI00000D3829; GB:AE001551; GB:AE001439; NID:g4155887; PIDN:
C:Genetics:
A:Gene: jhp1443
A:Note: similar to H. pylori 26695 gene HP1365
C:Superfamily: ompR protein; response regulator homology
C:Keywords: phosphoprotein
F:5-115/Domain: response regulator homology <RRH>
F:53/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 48.1%; Score 39; DB 2; Length 213;
Best Local Similarity 63.6%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SQEFLSHGGY 12
    |||||
Db 17 SIKEFLEHLGY 27

RESULT 26
T21527
hypothetical protein F28H7.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21527
R:Berks, M.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19435
A:Accession: T21527
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <WIL>
A:Cross-references: UNIPROT:Q19893; UNIPARC:UPI0000080FE7; EMBL:Z72508; PIDN:CAA96638.1.1
A:Experimental source: clone F28H7
C:Genetics:
A:Gene: CESP:F28H7.6
A:Map position: 5
A:Introns: 26/1; 94/2; 143/1; 175/1; 218/1; 325/1; 358/1; 401/1; 471/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F28H7.6

Query Match 48.1%; Score 39; DB 2; Length 525;
Best Local Similarity 42.9%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 CSIQEFLSHGGY 14
    |||||
Db 1 CSIQEFLSHGGY 14
```

```
Db      283 CSRQHIFEHNYMC 296

RESULT 27
C90154
primase SS00140 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: C90154
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-532 <KUR>
A;Cross-references: UNIPROT:Q980Y2; UNIPARC:UPI0000064196; GB:AE006641; NID:gl3813270; F
C;Genetics:
A;Gene: SS00140

Query Match      48.1%; Score 39; DB 2; Length 532;
Best Local Similarity 53.8%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      1 CSLQEFLSHGQYV 13
      |: :||| |
Db      93 CTLEWKISHGRYV 105

RESULT 28
T35002
probable respiratory chain oxidoreductase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35002
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21564
A;Accession: T35002
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-646 <SEE>
A;Cross-references: UNIPROT:Q9ZBV8; UNIPARC:UPI00000DAE34; EMBL:AL034443; PIDN:CAA22364.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC08DB:SC4B5.10C
C;Superfamily: hydrogen dehydrogenase alpha chain; NADH dehydrogenase (ubiquinone) I che

Query Match      48.1%; Score 39; DB 2; Length 646;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      2 SLQEFLSHGQY 12
      || :||| |
Db      257 SLDDYRAHGQY 267

RESULT 29
B83725
hypothetical protein BH0602 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: B83725
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83725
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-652 <STO>
A;Cross-references: UNIPROT:Q9KF83; UNIPARC:UPI000000C3921; GB:AP001509; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH0602

Query Match      48.1%; Score 39; DB 2; Length 652;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      3 LQBFSLHGQY 12
      | :||| |
Db      168 LQGFNLHGQY 177

RESULT 30
AG0265
alcohol dehydrogenase (EC 1.1.1.1) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 31-Dec-2004
C;Accession: AG0265
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, F.
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0265
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-891 <KUR>
A;Cross-references: UNIPROT:Q8ZEJ0; UNIPARC:UPI00000CD8B1; GB:AL590842; PIDN:CAC90987.1;
C;Genetics:
A;Gene: adhG
C;Superfamily: bifunctional aldehyde/alcohol dehydrogenase; aldehyde dehydrogenase homol
C;Keywords: oxidoreductase

Query Match      48.1%; Score 39; DB 2; Length 891;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      4 QEFSLHGQYV 13
      : :||| |
Db      264 ERFASHGGYL 273

RESULT 31
AE0650
alcohol dehydrogenase (EC 1.1.1.1) [imported] - Salmonella enterica subsp. enterica serov
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 31-Dec-2004
C;Accession: AE0650
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov.
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE0650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-892 <PAR>
A;Cross-references: UNIPARC:UPI0000059F55; GB:AL513382; PIDN:CAD08384.1; PID:gl6502428; F
C;Genetics:
A;Gene: adh
C;Superfamily: bifunctional aldehyde/alcohol dehydrogenase; aldehyde dehydrogenase homol
C;Keywords: oxidoreductase

Query Match      48.1%; Score 39; DB 2; Length 892;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C;Accession: A32519; S51854; S42150; S42151
R;Duncan, K.; Edwards, R.M.; Coggins, J.R.
Biochem. J. 246, 375-386, 1987
A;Title: The pentafunctional AROM enzyme of *Saccharomyces cerevisiae* is a mosaic of mono
A;Reference number: A32519; MUID:88076802; PMID:2825635
A;Accession: A32519
A;Molecule type: DNA
A;Residues: 1-1588 <DUN>
A;Cross-references: UNIPROT:P08566; UNIPARC:UPI0000052F13; EMBL:X06077; NID:g3380; PIDN:
R;Oliver, K.; Harris, D.
submitted to the EMBL Data Library, February 1995
A;Reference number: S51853
A;Accession: S51854
A;Molecule type: DNA
A;Residues: 1-1588 <OLI>
A;Cross-references: UNIPARC:UPI0000052F13; EMBL:Z48179; NID:9665657; PIDN:CAA88208.1; PI
R;Duncan, K.; Edwards, R.M.; Coggins, J.R.
FEBS Lett. 241, 83-86, 1988
A;Title: The *Saccharomyces cerevisiae* ARO1 gene. An example of the co-ordinate regulation
A;Reference number: S42150; MUID:89065111; PMID:2848727
A;Accession: S42150
A;Molecule type: DNA
A;Residues: 1-44;1557-1588 <DUN>
A;Cross-references: UNIPARC:UPI0000168B07; UNIPARC:UPI0000168B08; EMBL:X13802; EMBL:X138
C;Genetics:
A;Gene: SGD:ARO1; MIPS:YDR127W
A;Cross-references: SGD:S0002534; MIPS:YDR127W
A;Map position: 4R
C;Superfamily: pentafunctional AROM protein; 3-dehydroquinase dehydratase homology; 3-de
C;Keywords: aromatic amino acid biosynthesis; ATP; carbon-oxygen lyase; hydro-lyase; mul
F;7-380/Domain: 3-dehydroquinase synthase homology <DQS>
F;417-863/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>
F;887-1008/Domain: shikimate kinase homology <SKI>
F;1078-1291/Domain: 3-dehydroquinase dehydratase homology <DQD>
F;1362-1572/Domain: shikimate dehydrogenase homology <SKD>

Query Match 48.1%; Score 39; DB 1; Length 1588;

Best Local Similarity 58.3%; Pred. No. 2.9e+02;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYV 13

Db 981 ALKDPASGGYV 992

RESULT 34

T19130

hypothetical protein C09F9.2 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T19130

R;Smye, R.
submitted to the EMBL Data Library, November 1996

A;Reference number: Z19078

A;Accession: T19130

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: DNA

A;Residues: 1-1642 <WIL>

A;Cross-references: UNIPROT:O62055; UNIPARC:UPI000007E384; EMBL:Z81465; PIDN:CAB03861.1

A;Experimental source: clone C09F9

C;Genetics:

A;Gene: CESP:C09F9.2

A;Map position: 2

A;Introns: 75/3; 103/2; 220/1; 552/1; 1500/3; 1517/1; 1578/3

Query Match 48.1%; Score 39; DB 2; Length 1642;

Best Local Similarity 57.1%; Pred. No. 3e+02;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYV 14

|||||:|:|

QY 4 QEFSLSHGGYV 13

Db 264 ERFASHGGYV 273

RESULT 32

ATP-dependent permease ADP1 precursor - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein YCR011c; protein YCR105

C;Species: *Saccharomyces cerevisiae*

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C;Accession: S19421; S40914

R;Goffeau, A.; Purnelle, B.; Skala, J.
submitted to the Protein Sequence Database, March 1992

A;Reference number: S19420

A;Accession: S19421

A;Molecule type: DNA

A;Residues: 1-1049 <GOF>

A;Cross-references: UNIPROT:P25371; UNIPARC:UPI00001255FD; EMBL:X59720; NID:g1907116; PI

R;Purnelle, B.; Skala, J.; Goffeau, A.

Yeast 7, 867-872, 1991

A;Title: The product of the YCR105 gene located on the chromosome III from *Saccharomyces*

A;Reference number: S40914; MUID:92160395; PMID:1789009

A;Accession: S40914

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-1049 <PUR>

A;Cross-references: UNIPARC:UPI00001255FD

R;Skala, J.; Purnelle, B.; Goffeau, A.

Yeast 8, 409-417, 1992

A;Title: The complete sequence of a 10.8 kb segment distal of SUP2 on the right arm of C

K genes.

A;Reference number: S25353; MUID:92327849; PMID:1626432

A;Contents: annotation

C;Genetics:

A;Gene: SGD:ADP1; MIPS:YCR011C

A;Cross-references: SGD:S0000604; MIPS:YCR011C

A;Map position: 3R

C;Superfamily: ATP-dependent permease ADP1; ATP-binding cassette homology

C;Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein

F;1-25/Domain: signal sequence #status predicted <SIG>

F;126-1049/Product: ATP-dependent permease ADP1 #status predicted <MAT>

F;126-324/Domain: extracellular #status predicted <EXT>

F;325-341/Domain: transmembrane #status predicted <ABC>

F;406-607/Domain: ATP-binding cassette homology <ABC>

F;423-430/Region: nucleotide-binding motif A (P-loop)

F;550-557/Region: nucleotide-binding motif B

F;794-810/Domain: transmembrane #status predicted <TM2>

F;829-845/Domain: transmembrane #status predicted <TM3>

F;878-894/Domain: transmembrane #status predicted <TM4>

F;909-925/Domain: transmembrane #status predicted <TM5>

F;938-954/Domain: transmembrane #status predicted <TM6>

F;1025-1041/Domain: transmembrane #status predicted <TM7>

F;50,114,165,221/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;429/Binding site: ATP (Lys) #status predicted

Query Match 48.1%; Score 39; DB 1; Length 1049;

Best Local Similarity 50.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LOEFSLSHGGYV 14

Db 618 VSEFLRNEGYIC 629

RESULT 33

BVBYA1

3-dehydroquinase synthase (EC 4.2.3.4) - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: protein YD9302.02; protein YDR127W

N;Contents: 3-dehydroquinase dehydratase (EC 4.2.1.10); 3-dehydroquinase synthase (EC 4.

ikimate kinase (EC 2.7.1.71)

C;Species: *Saccharomyces cerevisiae*

Db 494 CSLNETLNECGRVC 507

RESULT 35

E86189

hypothetical protein [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: E86189

R/TheoLogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E86189

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1878 <STO>

A/Cross-references: UNIPROT:Q9SVJ7; UNIPARC:UPI000009DC9E; GB:AE005172; NID:g4836907; PID

C/Genetics:

A/Map position: 1

Query Match 47.5%; Score 38.5; DB 2; Length 1878;

Best Local Similarity 69.2%; Pred. No. 4.1e+02;

Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 3 LQEFLLSHGGYVC 14

||||| |||

Db 1300 LQEFLEKHGGYVC 1312

RESULT 36

T03751

hypothetical protein A - slime mold (Physarum polycephalum) mitochondrion integrated pla

C/Species: mitochondrion Physarum polycephalum

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C/Accession: T03751

R/Nakagawa, C.C.; Jones, E.P.; Miller, D.L.

Curr. Genet. 33, 178-187, 1998

A/Title: Mitochondrial DNA rearrangements associated with mF plasmid integration and pla

A/Reference number: Z15055; MUID:98177147; PMID:9508792

A/Accession: T03751

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-194 <NAK>

A/Cross-references: UNIPROT:O20627; UNIPARC:UPI000008B89C; EMBL:AF012249; NID:g2435419;

A/Experimental source: strain aux2-S

C/Genetics:

A/Genome: mitochondrion integrated plasmid mF

C/Keywords: mitochondrion

Query Match 46.9%; Score 38; DB 2; Length 94;

Best Local Similarity 42.9%; Pred. No. 24;

Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLQEFLLSHGGYVC 14

||| : ||| |||

Db 77 CYIDKPCSIMGYIC 90

RESULT 37

A82985

hypothetical protein PAS285 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C/Accession: A82985

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: A82985

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-105 <STO>

A/Cross-references: UNIPROT:Q9HTR9; UNIPARC:UPI000000CF89; GB:AE004941; GB:AE004091; NID

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PAS285

Query Match 46.9%; Score 38; DB 2; Length 105;

Best Local Similarity 58.3%; Pred. No. 27;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFLLSHGGYVC 13

||||| |||

Db 70 AMEEFLSRGGKV 81

RESULT 38

S46445

YMP46 protein - Acanthamoeba castellanii mitochondrion

C/Species: mitochondrion Acanthamoeba castellanii

C/Date: 26-Dec-1994 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004

C/Accession: S46445; S53825

R/Loneragan, K.M.; Gray, M.W.

J. Mol. Biol. 239, 476-499, 1994

A/Title: The ribosomal RNA gene region in Acanthamoeba castellanii mitochondrial DNA. A

A/Reference number: S46445; MUID:94275838; PMID:8006963

A/Accession: S46445

A/Molecule type: DNA

A/Residues: 1-142 <LON>

A/Cross-references: UNIPROT:Q37060; UNIPARC:UPI000008BB47; EMBL:U03732; NID:g495327; PID

R/Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.

J. Mol. Biol. 245, 522-537, 1995

A/Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: comp

A/Reference number: S53825; MUID:95147275; PMID:7844823

A/Accession: S53825

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-142 <BUR>

A/Cross-references: UNIPARC:UPI000008BB47; GB:U12386; NID:g562028; PIDN:AAD1817.1; PID;

A/Experimental source: strain Neff; ATCC 30010

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C/Genetics:

A/Genome: mitochondrion

A/Genetic code: SGC6

C/Keywords: mitochondrion

Query Match 46.9%; Score 38; DB 2; Length 142;

Best Local Similarity 38.5%; Pred. No. 37;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SLQEFLLSHGGYVC 14

||||| |||

Db 52 AIKDFFGHGSVVC 64

RESULT 39

JE0136

lectin precursor - common snowdrop

C/Species: Galanthus nivalis (common snowdrop)

C/Date: 02-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 31-Dec-2004

C/Accession: JE0136; JCS5675

R/Zhou, Y.; Tian, Y.; Tao, H.; Mang, K.

Chinese J. Biotechnol. 12, 495-498, 1996

A/Title: Molecular cloning and sequence analysis of snowdrop (Galanthus nivalis) lectin

A/Reference number: JE0136

A/Accession: JE0136

A:Molecule type: DNA
A:Residues: 1-160 <ZHO>
A:Cross-references: UNIPARC:UPI0000179130
A:Accession: J05675
A:Molecule type: DNA
A:Residues: 1-160 <ZH2>
A:Cross-references: UNIPARC:UPI0000179130
A:Experimental source: leaf
C:Superfamily: lectin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:27-131/Product: lectin #status predicted <WAT>
F:132-160/Domain: carboxyl-terminal cleaved off #status predicted <CTC>

Query Match 46.9%; Score 38; DB 2; Length 160;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYYV 13
|:|||||:
DB 37 STGEFLNYGGFV 48
|:|||||:

RESULT 40
E82780
response regulator - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: E84690
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L. Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: E84690
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-213 <TOM>
A:Cross-references: UNIPROT:Q25918; UNIPARC:UPI000003157; GB:AE000636; GB:AE000511; NID
C:Superfamily: ompR protein; response regulator homology
C:Keywords: phosphoprotein
F:5-115/Domain: response regulator homology <RRH>
F:53/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 46.9%; Score 38; DB 2; Length 213;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SLQEFLSHGYY 12
|:|||||:
DB 17 SVKEFLEHLGY 27
|:|||||:

RESULT 41
E82780
endonuclease III XF0647 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82780
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82780
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <SIM>
A:Cross-references: UNIPROT:Q9PFL0; UNIPARC:UPI00000C24A9; GB:AE003909; GB:AE003849; NID
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A. Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, D.M.; Carrer, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.B.; Kuramae, E.E.; Laigr Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0647
C:Superfamily: apurinic/apyrimidinic endonuclease III
C:Keywords: 4Fe-4S; metalloprotein
F:193,200,203,209/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 46.9%; Score 38; DB 2; Length 218;
Best Local Similarity 47.4%; Pred. No. 57;
Matches 9; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 4 QEFLS-----HGGYVC 14
|:|||||:
DB 175 QEFKDAHHWILHGRYVC 193
|:|||||:

RESULT 42
D75275
endonuclease III - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75275
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.W. Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75275
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <WHI>
A:Cross-references: UNIPROT:Q9RRQ0; UNIPARC:UPI00000D3FB7; GB:AE002073; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2438
A:Map position: 1

Query Match 46.9%; Score 38; DB 2; Length 259;
Best Local Similarity 77.8%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 FLSHGGYVC 14
|:|||||:
DB 209 FLSHGRQVC 217
|:|||||:

RESULT 43
AH2340
transcription regulator all4279 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Title: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2340
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpso, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium An A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2340
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-295 <KUR>
A;Cross-references: UNIPROT:Q8VPB9; UNIPARC:UPI00000CEA96; GB:BA000019; PIDN:BAB75978.1;
A;Experimental source: strain FCC 7120
C;Genetics:
A;Gene: all4279

Query Match 46.9%; Score 38; DB 2; Length 295;
Best Local Similarity 66.7%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHG 10
:|||||:
Db 29 AMQEFLLTHG 37

RESULT 44
T01266
starch synthase DULL1 - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01266
R;Gao, M.; Wanat, J.; Stinard, P.S.; James, M.G.; Myers, A.M.
Plant Cell 10, 399-412, 1998
A;Title: Characterization of dull1, a maize gene coding for a novel starch synthase.
A;Reference number: 214287; MUID:98169346; PMID:9501113
A;Accession: T01266
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-310 <GAO>
A;Cross-references: UNIPROT:O64924; UNIPARC:UPI00000A655C; EMBL:AF023160; NID:G3057121;
C;Genetics:
A;Gene: dull1
A;Map position: 10
A;Note: Intron positions not resolved (incomplete sequence)

Query Match 46.9%; Score 38; DB 2; Length 310;
Best Local Similarity 61.5%; Pred. No. 81;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHGGYVC 14
||| |||
Db 294 SLRMLSSGMVVC 306

RESULT 45
D83708
phosphonates transport system (phosphate-binding protein) BH0468 [imported] - Bacillus h
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: D83708
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83708
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-323 <STO>
A;Cross-references: UNIPROT:Q9KFL1; UNIPARC:UPI00000D73BA; GB:AP001508; GB:BA0000004; NID
C;Genetics:
A;Gene: BH0468

Query Match 46.9%; Score 38; DB 2; Length 323;
Best Local Similarity 63.6%; Pred. No. 85;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLQEFSLSHGGY 12
|:|||||
Db 282 SMQEFLESNGY 292

RESULT 46

E71363
hypothetical protein TP0118 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: E71363
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
rvey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: E71363
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-421 <COL>
A;Cross-references: UNIPROT:O83155; UNIPARC:UPI00001392CC; GB:AE001197; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0118

Query Match 46.9%; Score 38; DB 2; Length 421;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CSLQEFSLSHGGYVC 14
||| | : |||
Db 26 CSLGEHIVRGGAAC 39

RESULT 47
E64657
Probable glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) - Helicobacter pylori (strain
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: E64657
R;Tombr, J.P.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Petersen, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.M.; Venter, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64657
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-425 <TOM>
A;Cross-references: UNIPROT:P56110; UNIPARC:UPI000012AF23; GB:AE000616; GB:AE000511; NID
C;Superfamily: Glucose-6-phosphate dehydrogenase
C;Keywords: oxidoreductase

Query Match 46.9%; Score 38; DB 2; Length 425;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 OEFLSHGGYVC 14
:||||| :|:
Db 69 REFLAHISYLC 79

RESULT 48
A33504
Glutamate dehydrogenase (NADP) (EC 1.4.1.4) - Salmonella typhimurium
N;Alternate names: glutamic dehydrogenase; NADP-specific glutamate dehydrogenase
C;Species: Salmonella typhimurium
C;Date: 08-Dec-1989 #sequence_revision 23-Mar-1995 #text_change 05-Oct-2004
C;Accession: A33504
R;Bansal, A.; Dayton, M.A.; Zalkin, H.; Colman, R.F.
J. Biol. Chem. 264, 9827-9835, 1989
A;Title: Affinity labeling of a glutamyl peptide in the coenzyme binding site of NADP(+)
enosine 2',5'-bisphosphate.
A;Reference number: A33504; MUID:89255551; PMID:2656714
A;Accession: A33504
A;Molecule type: DNA

A;Residues: 1-447 <BAN>
A;Cross-references: UNIPROT:P15111; UNIPARC:UPI0000129306; GB:M24021; GB:J04814
C;Genetics:
A;Gene: gdh
A;Map position: 27 min
C;Superfamily: glutamate dehydrogenase
C;Keywords: homohexamer; NADP; oxidoreductase
F;128/Binding site: substrate (lys) #status predicted

Query Match 46.9%; Score 38; DB 1; Length 447;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSLOEFLSH 9
|||: ||:|
Db 5 CSLESFLNH 13

RESULT 49
AF0710
NADP-specific glutamate dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typhi
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004
C;Accession: AF0710
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Aitken, S.; Brown, J.E.; et al.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0710
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <PAR>
A;Cross-references: UNIPARC:UPI0000059DA7; GB:AL513382; PIDN:CAD02055.1; PID:gi6502892;
C;Genetics:
A;Gene: STY1815
C;Superfamily: glutamate dehydrogenase

Query Match 46.9%; Score 38; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSLOEFLSH 9
|||: ||:|
Db 5 CSLESFLNH 13

RESULT 50
E83130
conserved hypothetical protein PA4132 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83130
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Brinkman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: AB2950; MUID:2043737; PMID:10984043
A;Accession: E83130
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <STO>
A;Cross-references: UNIPROT:Q9HWP9; UNIPARC:UPI000005C529; GB:AE004829; GB:AE004091; NID:1000000000
C;Genetics:
A;Experimental source: strain PA01
A;Gene: PA4132

Query Match 46.9%; Score 38; DB 2; Length 471;
Best Local Similarity 45.5%; Pred. No. 1.2e+02;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGY 12

Db 352 AIADYLOHGGY 362

Search completed: May 12, 2006, 10:52:30
Job time : 26.0915 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds
(without alignments)

99.694 Million cell updates/sec

Title: US-10-632-388-53

Perfect score: 81

Sequence: 1 CSLQFLSHGCVVC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	60.5	488	2	Q6CNV5_KLUJLA
2	48	59.3	362	2	Q5S4N1_HUMAN
3	48	59.3	365	2	Q9FNE6_ARATH
4	48	59.3	367	2	Q8GUG6_ARATH
5	48	59.3	480	2	Q4R9A4_MACFA
6	48	59.3	658	1	Q4R9A4_MACFA
7	48	59.3	660	1	Q4R9A4_MACFA
8	48	59.3	660	1	Q4R9A4_MACFA
9	48	59.3	660	1	Q4R9A4_MACFA
10	48	59.3	661	1	Q4R9A4_MACFA
11	48	59.3	661	1	Q4R9A4_MACFA
12	48	59.3	662	1	Q4R9A4_MACFA
13	48	59.3	662	1	Q4R9A4_MACFA
14	48	59.3	674	2	Q5GKX6_HUMAN
15	46	56.8	202	2	Q8QH15_TYRTY
16	46	56.8	358	2	Q7XK97_ORYSA
17	46	56.8	358	2	Q69BLO_MANSE
18	45	55.6	363	2	Q932D2_ARATH
19	45	55.6	363	2	Q9C5A2_ARATH
20	45	55.6	376	2	Q5ER66_MONAL
21	45	55.6	448	2	Q518T7_ENTHI
22	45	55.6	570	1	RS1_CHILMU
23	45	55.6	611	2	Q4KMN1_MOUSE
24	45	55.6	651	2	Q5F491_CHICK
25	45	55.6	672	2	Q8BIR2_MOUSE
26	45	55.6	695	1	Q52L23_XENLA
27	45	55.6	697	1	AN3_XENLA
28	45	55.6	697	2	Q7ZXJ0_XENLA
29	45	55.6	699	2	Q6P4J3_XENTRO
30	45	55.6	948	2	Q51FR1_ENTHI
31	44.5	54.9	651	2	Q9FUE2_ARATH

32	44	54.3	266	2	Q5DELO_SCHJA
33	44	54.3	319	2	Q81633_ARATH
34	44	54.3	319	2	Q9SBA5_ARATH
35	44	54.3	342	2	Q7XBW0_ORYSA
36	44	54.3	342	2	Q84Y01_MAIZE
37	44	54.3	342	2	Q8W3H7_ORYSA
38	44	54.3	348	2	Q92M44_RHIME
39	44	54.3	388	2	Q98S45_RHIL0
40	44	54.3	393	2	Q95K70_MACFA
41	44	54.3	413	2	Q6AJD7_DESFS
42	44	54.3	429	1	MYCS_RAT
43	44	54.3	430	2	Q7N4F2_PHOLL
44	44	54.3	489	2	Q5R8M2_PONPY
45	44	54.3	490	2	Q6UX62_HUMAN
46	44	54.3	490	2	Q5VX71_HUMAN
47	44	54.3	701	2	Q4L8P7_STAHR
48	44	54.3	123	2	Q07363_CHLTR
49	43	53.1	219	2	Q4RUI2_TETNG
50	43	53.1	469	2	Q759D9_ASHGO
51	43	53.1	490	2	Q8BH32_MOUSE
52	43	53.1	492	2	Q8VC43_MOUSE
53	43	53.1	516	2	Q608U7_METCA
54	43	53.1	521	2	Q7PYM7_ANOGA
55	43	53.1	666	2	Q5LXF8_SILPO
56	42	51.9	93	2	Q7N3Y6_GLOVI
57	42	51.9	292	2	Q60ML3_CABBR
58	42	51.9	293	2	Q18639_CAEEL
59	42	51.9	308	2	Q88J21_PSEPK
60	42	51.9	333	2	Q7XHG0_ORYSA
61	42	51.9	333	2	Q9AYJ2_ORYSA
62	42	51.9	338	2	Q9SUG3_ARATH
63	42	51.9	353	2	Q93VQ8_ARATH
64	42	51.9	355	2	Q6K7B8_ORYSA
65	42	51.9	357	2	Q75GI4_ORYSA
66	42	51.9	357	2	Q8GS15_ORYSA
67	42	51.9	439	1	GLNA_THENA
68	42	51.9	615	2	Q7V9F0_PROMA
69	42	51.9	626	2	Q7U3A8_SYNPM
70	42	51.9	629	2	Q7V3R9_PRONM
71	42	51.9	789	2	Q8TXR3_METKA
72	42	51.9	1871	2	Q67FT1_MESMA
73	41.5	51.2	113	2	Q8X210_ECO57
74	41.5	51.2	113	2	Q6H9T6_9CAUD
75	41.5	51.2	2043	2	Q54C50_DICTDI
76	41	50.6	100	2	Q9PLH9_HUMAN
77	41	50.6	143	2	Q56C40_9CAUD
78	41	50.6	194	2	Q919G7_9YVRA
79	41	50.6	216	2	Q804U8_SAYPH
80	41	50.6	274	2	Q5XIK0_RAT
81	41	50.6	280	2	Q8STN6_ENCCU
82	41	50.6	320	2	Q88H11_PSEPK
83	41	50.6	352	2	Q7P029_CHRVO
84	41	50.6	354	2	Q8BJX2_MOUSE
85	41	50.6	363	2	Q4QL17_LEIMA
86	41	50.6	423	2	Q5CUE9_CRYPV
87	41	50.6	436	2	Q9XU02_CAEEL
88	41	50.6	470	1	TRAF4_HUMAN
89	41	50.6	470	1	Q8BHD3_MOUSE
90	41	50.6	470	2	Q6DFG9_XENLA
91	41	50.6	470	2	Q6DEV2_XENTR
92	41	50.6	470	2	Q68F27_XENLA
93	41	50.6	470	2	Q7T1K2_BRARE
94	41	50.6	471	2	Q4E8L8_TETNG
95	41	50.6	503	2	Q75EB7_ASHGO
96	41	50.6	518	2	Q89QJ1_BRAJA
97	41	50.6	602	2	Q5IT31_MAGGR
98	41	50.6	873	2	Q64Y89_BACFR
99	41	50.6	888	2	Q5LHH5_BACFR
100	41	50.6	1874	2	Q627X6_CAEEL
101	41	50.6	2395	2	Q6Q9H2_CAEEL
102	41	50.6	2407	2	Q9NEL2_CAEEL
103	41	50.6	2407	2	Q9NEL2_CAEEL
104	40.5	50.0	87	2	Q515N5_ENTHI

Q5del0	schistosoma
Q81633	arabidopsis
Q9SBA5	arabidopsis
Q7XBW0	oryza sativ
Q84Y01	zea mays (m
Q8W3H7	oryza sativ
Q92M44	rhizobium m
Q98S45	rhizobium l
Q95K70	macaca fasc
Q6AJD7	desulfotale
P23999	rattus norv
Q7N4F2	photorhabdu
Q5r8m2	pongo pygma
Q5ux62	homo sapien
Q5vx71	homo sapien
Q4L8P7	staphylococ
Q07363	chlamydia t
Q4RUI2	tetradon n
Q759D9	ashbya goss
Q8BH32	m mus muscu
Q8VC43	mus musculu
Q608U7	methylococc
Q7PYM7	anopheles g
Q5LXF8	silicibacte
Q7N3Y6	gloeobacter
Q60ML3	caenorhabdi
Q18639	caenorhabdi
Q88J21	pseudomonas
Q7XHG0	oryza sativ
Q9AYJ2	oryza sativ
Q9SUG3	arabidopsis
Q93VQ8	arabidopsis
Q6K7B8	oryza sativ
Q75GI4	oryza sativ
Q8GS15	oryza sativ
P3205	thermotoga
Q7V9F0	prochloroco
Q7U3A8	synecchococ
Q7V3R9	prochloroco
Q8TXR3	methanopyru
Q67FT1	mesobutrus
Q8X210	escherichia
Q6H9T6	phage phi 4
Q54C50	dictyosteli
Q9PLH9	homo sapien
Q56C40	enterobacte
Q919G7	elaenia fla
Q804U8	sayornis ph
Q5XIK0	rattus norv
Q8STN6	encephalito
Q88H11	pseudomonas
Q7P029	chromobacte
Q8BJX2	mus musculu
Q4QL17	leishmania
Q5CUE9	cryptospori
Q9XU02	caenorhabdi
Q6Buz4	homo sapien
Q61382	mus musculu
Q8BHD9	m mus muscu
Q6DFG9	xenopus lae
Q6DEV2	xenopus tro
Q68F27	xenopus lae
Q7T1K2	brachydanio
Q4E8L8	tetradon n
Q75EB7	ashbya goss
Q89QJ1	bradyrhizob
Q5IT31	magnaporthe
Q64Y89	bacteroides
Q5LHH5	bacteroides
Q627X6	caenorhabdi
Q6Q9H2	caenorhabdi
Q9NEL2	caenorhabdi
Q515N5	entamoeba h

105	40.5	50.0	719	2	Q86S17_CABEL	Q86S17 caenorhabdi	178	39	48.1	423	2	Q5CGN4_CRYHO	Q5CGN4 cryptospori
106	40.5	50.0	733	2	Q4QPY0_LEIMA	Q4qfy0 leishmania	179	39	48.1	437	2	Q5C307_SCHJA	Q5c307 schistosoma
107	40.5	50.0	880	2	Q9TXM6_CABEL	Q9txm6 caenorhabdi	180	39	48.1	510	1	CP191_CARAU	P79690 carassius a
108	40.5	50.0	936	2	Q624X9_CABER	Q624x9 caenorhabdi	181	39	48.1	512	2	Q88XS6_LACPL	Q88xs6 lactobacill
109	40	49.4	179	2	Q67B68_SCHLR	Q67b68 dehalococco	182	39	48.1	513	2	Q5FKK0_LACAC	Q5fkko lactobacill
110	40	49.4	184	2	Q5ZUT8_LEGPH	Q5zut8 legionella	183	39	48.1	525	2	Q19893_CABEL	Q19893 caenorhabdi
111	40	49.4	189	2	Q6VS08_9PASS	Q6vsq8 muscicapa a	184	39	48.1	531	2	Q60YX3_CABER	Q60yx3 caenorhabdi
112	40	49.4	243	2	Q7Z3J8_HUMAN	Q7z3j8 homo sapien	185	39	48.1	532	2	Q980V2_SULSO	Q980v2 sulfolobus
113	40	49.4	271	2	Q870R8_NEUCR	Q870p8 neurospora	186	39	48.1	568	2	Q6E437_CUCME	Q6e437 cucumis mel
114	40	49.4	272	2	Q6Z9U1_BURMA	Q6z9u1 burkholderi	187	39	48.1	607	2	Q82M25_STRAW	Q82m25 streptomyc
115	40	49.4	284	2	Q7RZJ6_NEUCR	Q7rzj6 neurospora	188	39	48.1	609	2	Q9GLF0_CANFA	Q9glf0 canis famil
116	40	49.4	294	2	Q8TI08_METAC	Q8ti08 methanosaec	189	39	48.1	611	1	LYAM2_CANFA	P33730 canis famil
117	40	49.4	350	2	Q4NBW3_WMICC	Q4nbw3 arthrobacte	190	39	48.1	618	2	Q7UZF4_PROMP	Q7uzf4 prochloroc
118	40	49.4	375	2	Q9XC94_KLEPN	Q9xc94 klebsiella	191	39	48.1	646	2	Q9ZBV8_STRCO	Q9zbv8 streptomyc
119	40	49.4	375	2	Q4QTM7_KLEPN	Q4qtm7 klebsiella	192	39	48.1	652	2	Q9K983_BACHD	Q9kf83 bacillus ha
120	40	49.4	380	2	Q7DMS9_TOBAC	Q7dms9 nicotiana t	193	39	48.1	661	2	Q8K043_MOUSE	Q8k043 mus musculu
121	40	49.4	392	2	Q43326_TOBAC	Q43326 nicotiana t	194	39	48.1	728	2	Q5RIY5_MALDO	Q5rly5 malus domes
122	40	49.4	436	2	Q4YRZ7_PLABE	Q4yrz7 plasmodium	195	39	48.1	748	1	SORT_RAT	Q54861 rattus norv
123	40	49.4	461	1	SSL1_YEAST	Q04673 saccharomyc	196	39	48.1	812	2	Q7S5T9_NEUCR	Q7s5t9 neurospora
124	40	49.4	461	2	Q8B27_YEAST	Q8b237 saccharomyc	197	39	48.1	825	2	Q9QXW6_MOUSE	Q9qxw6 mus musculu
125	40	49.4	472	2	Q4ZRV3_FSESY	Q4zrv3 pseudomonas	198	39	48.1	825	2	Q6PHU5_MOUSE	Q6phus mus musculu
126	40	49.4	478	2	Q7TLK1_BRARE	Q7tlk1 brachydanio	199	39	48.1	827	2	Q5CKM9_CRYHO	Q5ckm9 cryptospori
127	40	49.4	497	1	ARAA_BACHD	Q9kbg2 bacillus ha	200	39	48.1	831	1	SORT_HUMAN	Q99523 homo sapien
128	40	49.4	500	1	ATG4_PODAN	Q86z15 podospora a	201	39	48.1	831	2	Q8I249_HUMAN	Q8i249 homo sapien
129	40	49.4	527	2	Q5WBZ7_BACSK	Q5wbz7 bacillus cl	202	39	48.1	833	2	Q8N6W5_HUMAN	Q8n6w5 homo sapien
130	40	49.4	562	2	Q869B7_LEIME	Q869b7 leishmania	203	39	48.1	834	2	Q5CVI3_CRYPV	Q5cvl3 cryptospori
131	40	49.4	562	2	Q8I911_LEIAM	Q8i911 leishmania	204	39	48.1	891	2	Q8D0K2_YERPE	Q8d0k2 yersinia pe
132	40	49.4	562	2	Q4Q3K1_LEIMA	Q4q3k1 leishmania	205	39	48.1	891	2	Q8ZEU0_YERPE	Q8zej0 yersinia pe
133	40	49.4	569	1	RS1_CHLTR	Q84100 chlamydia t	206	39	48.1	891	2	Q66AM7_YERPS	Q66am7 yersinia ps
134	40	49.4	577	2	Q5SY54_HUMAN	Q5sy54 homo sapien	207	39	48.1	892	2	Q57NR1_SALCH	Q57nr1 salmonella
135	40	49.4	672	2	Q4H484_9DEIO	Q4h484 deinococcus	208	39	48.1	892	2	Q8Z7F1_SALTY	Q8z7f1 salmonella
136	40	49.4	688	2	Q4T0K3_TETNG	Q4t0k3 tetradodon n	209	39	48.1	892	2	Q8ZP45_SALTY	Q8zpz45 salmonella
137	40	49.4	802	2	Q60YM9_CABER	Q60ym9 caenorhabdi	210	39	48.1	892	2	Q5PCV6_SALPA	Q5pcv6 salmonella
138	40	49.4	802	2	Q5YQ04_CABEL	Q5yqc4 caenorhabdi	211	39	48.1	908	2	Q9C0G5_HUMAN	Q9c0g5 homo sapien
139	40	49.4	890	2	Q5SY53_HUMAN	Q5sy53 homo sapien	212	39	48.1	973	2	Q5VR95_ORYSA	Q5vr95 oryza sativ
140	40	49.4	890	2	Q7Z616_HUMAN	Q7z616 homo sapien	213	39	48.1	979	2	Q4RPC8_TETNG	Q4rpc8 tetradodon n
141	40	49.4	917	2	Q80389_BRARE	Q80389 brachydanio	214	39	48.1	1019	2	Q6P9B9_HUMAN	Q6p9b9 homo sapien
142	40	49.4	1029	2	Q7RFV6_PLAYO	Q7rfv6 plasmodium	215	39	48.1	1036	2	Q8A3G5_BACTN	Q8a2g5 bacteroides
143	40	49.4	1084	2	Q5IVB3_MAGGR	Q5ivb3 magnaporth	216	39	48.1	1049	1	ADP_YEAST	P25371 saccharomyc
144	40	49.4	1093	1	SYNP2_HUMAN	Q9ums6 homo sapien	217	39	48.1	1103	2	Q54R25_DICDI	Q54r25 dictyosteli
145	40	49.4	1110	2	Q6ZML6_HUMAN	Q6zml6 homo sapien	218	39	48.1	1190	2	Q4KG87_PSEFS	Q4kg87 pseudomonas
146	40	49.4	1209	2	Q54SA2_DICDI	Q54sa2 dictyosteli	219	39	48.1	1234	2	Q5ZM67_CHICK	Q5zm67 gallus gall
147	40	49.4	1382	2	Q518V2_BRARE	Q518v2 brachydanio	220	39	48.1	1280	2	Q6QHS1_LYTVA	Q6qhs1 lytechinus
148	40	49.4	1520	2	Q4Q7T0_LEIMA	Q4q7t0 leishmania	221	39	48.1	1588	1	AR01_YEAST	P08566 a pentafunc
149	40	49.4	3283	2	Q4U1Z1_THEAN	Q4u1z1 theileria a	222	39	48.1	1638	2	Q61WT5_CABER	Q61wt5 caenorhabdi
150	39.5	48.8	113	2	Q8X3G6_ECO57	Q8x3g6 escherichia	223	39	48.1	1642	2	Q6ZWS1_TETNG	Q6zws1 tetradodon n
151	39.5	48.8	113	2	Q8X5F5_ECO57	Q8x5f5 escherichia	224	38.5	47.5	512	2	O4SWS1_TETNG	Q4sws1 tetradodon n
152	39.5	48.8	696	2	Q5CVJ1_CRYPV	Q5cvj1 cryptospori	225	38.5	47.5	665	1	KUP2_LACAC	Q5flf5 lactobacill
153	39	48.1	120	2	Q736R5_BACCI	Q736r5 bacillus ce	226	38.5	47.5	671	1	KUP2_LACAC	Q74ln2 lactobacill
154	39	48.1	134	2	Q82UJ9_NITEU	Q82uuj9 nitrosomona	227	38.5	47.5	746	2	Q4R845_TETNG	Q4re45 tetradodon n
155	39	48.1	139	2	Q9PYE2_9BROM	Q9pye2 cucumber mo	228	38.5	47.5	859	2	Q81FV3_BACCR	Q81fv3 bacillus ce
156	39	48.1	161	2	Q92X31_RHIME	Q92x31 rhizobium m	229	38.5	47.5	861	2	Q61187_BACAN	Q61187 bacillus an
157	39	48.1	165	2	Q9TNM0_METAC	Q9tnm0 methanosaec	230	38.5	47.5	861	2	Q4MK61_BACCE	Q4mk61 bacillus ce
158	39	48.1	187	2	Q9RY85_DEIRA	Q9ry85 deinococcus	231	38.5	47.5	861	2	Q6HL89_BACHK	Q6hl89 bacillus th
159	39	48.1	200	2	P87426_GAVES	P87426 gavia arcti	232	38.5	47.5	861	2	Q73B30_BACCI	Q73b30 bacillus ce
160	39	48.1	202	2	Q8QH35_9PASS	Q8qh35 lepidothrix	233	38.5	47.5	861	2	Q81T06_BACAN	Q81t06 bacillus an
161	39	48.1	213	2	Q9Z371_HELPU	Q9z371 helicobacte	234	38.5	47.5	861	2	Q63DR6_BACCC	Q63dr6 bacillus ce
162	39	48.1	226	2	Q4NI67_9BROM	Q4ni67 arthrobacte	235	38.5	47.5	901	2	Q9LR43_ARATH	Q9lr43 arabidopsais
163	39	48.1	230	2	Q93267_9BROM	Q93267 cucumber mo	236	38.5	47.5	1878	2	Q9SYJ7_ARATH	Q9syj7 arabidopsais
164	39	48.1	244	2	Q91BB7_NPVST	Q91bb7 spodoptera	237	38.5	47.5	1950	2	Q9AUE0_ARATH	Q9aue0 arabidopsais
165	39	48.1	283	2	Q8SMU0_ENCVU	Q8smu0 entomoeba h	238	38.5	47.5	2109	2	Q5R9F7_PONPV	Q5r9f7 pongo pygma
166	39	48.1	287	2	Q51E80_ENTHI	Q51e80 entomoeba h	239	38.5	47.5	8515	2	Q4WVN4_ASPTU	Q4wvn4 aspergillus
167	39	48.1	292	2	Q5CKF5_CRYHO	Q5ckf5 cryptospori	240	38	46.9	69	2	Q4WVN2_ASPTU	Q4wdn2 aspergillus
168	39	48.1	349	2	Q4YU8_TETNG	Q4ryu8 tetradodon n	241	38	46.9	71	2	Q4LINO_MICAG	Q4lino microtus ag
169	39	48.1	354	1	ALR_FUSN	Q9kfu5 tetradodon n	242	38	46.9	94	2	O2O627_PHYPO	Q2o627 physarum po
170	39	48.1	361	2	Q5AXD6_EMENI	Q5axd6 fusobacteri	243	38	46.9	105	2	Q9HTR9_PSEAE	Q9htr9 pseudomonas
171	39	48.1	399	2	Q7T0P4_XENLA	Q7t0p4 xenopus lae	244	38	46.9	109	2	Q8ACE9_PSEPK	Q8ace9 pseudomonas
172	39	48.1	406	2	Q50XJ2_ENTHI	Q50xj2 entomoeba h	245	38	46.9	117	2	Q84258_PAPVE	Q84258 european el
173	39	48.1	409	2	Q5SP67_CRYNE	Q5sp67 cryptococcu	246	38	46.9	122	2	Q4YNP2_PLABE	Q4ynp2 plasmodium
174	39	48.1	409	2	Q5KE90_CRYNE	Q5ke90 cryptococcu	247	38	46.9	126	2	Q4RBS7_TETNG	Q4rbs7 tetradodon n
175	39	48.1	416	2	Q39Z35_9GAMA	Q39z35 rheus monk	248	38	46.9	136	2	Q752L8_ASHGO	Q752l8 ashbya gos
176	39	48.1	416	2	Q77V18_9GAMA	Q77v18 macaca mula	249	38	46.9	142	2	Q37060_ACACA	Q37060 acanthamoeb
177	39	48.1	416	2	Q53D52_9GAMA	Q53d52 macaca fusc	250	38	46.9	145	2	Q4RPI8_TETNG	Q4rpi8 tetradodon n

251	38	46.9	166	2	Q6ZL45_ORYSA	Q6ZL45_oryza sativ	324	38	46.9	775	2	Q8DJ10_SYNEL	Q8dj10_synechococ
252	38	46.9	177	2	Q6Z6B5_CAEBR	Q6Z6b5_caenorhabdi	325	38	46.9	840	2	Q5THT2_HUMAN	Q5tht2_homo sapien
253	38	46.9	177	2	Q8BW2C_MOUSE	Q8bw2c_mus muscucu	326	38	46.9	867	2	Q4PG03_USTWA	Q4pg03_ustilago ma
254	38	46.9	198	2	Q4J5N8_AZOVI	Q4j5n8_azotobacter	327	38	46.9	877	2	Q7W5Z8_BORPA	Q7w5z8_bordetella
255	38	46.9	206	2	Q6MPH5_BDEBA	Q6mph5_bdellovibri	328	38	46.9	877	2	Q7WGU1_BORBR	Q7wgu1_bordetella
256	38	46.9	207	2	Q5LA87_BACFN	Q5la87_bacteroides	329	38	46.9	893	2	Q7VZG5_BORPE	Q7vzg5_bordetella
257	38	46.9	212	2	Q87BD3_XYLFT	Q87bd3_xylella fas	330	38	46.9	952	2	Q6CEB7_YARLI	Q6ceb7_yarrowia li
258	38	46.9	213	2	Q25918_HELPY	Q25918_helicobacte	331	38	46.9	961	2	Q9SXA3_ARATH	Q9sxa3_arabidopsis
259	38	46.9	215	1	END3_BUCBP	Q89aw4_buchera ap	332	38	46.9	1032	2	Q7YYX2_CRYPV	Q7yyx2_cryptospori
260	38	46.9	218	2	Q9PFL0_XYLFA	Q9pfl0_xylella fas	333	38	46.9	1081	2	Q7YUL8_9TRYP	Q7yul8_trypanosoma
261	38	46.9	225	2	Q5WM04_LEGPA	Q5wm04_legionella	334	38	46.9	1243	2	Q6ZQ05_HUMAN	Q6zq05_homo sapien
262	38	46.9	225	2	Q5X4L4_LEGPA	Q5x4l4_legionella	335	38	46.9	1672	2	Q5T4S8_HUMAN	Q5t4s8_homo sapien
263	38	46.9	225	2	Q5ZUV5_LEGPH	Q5zuv5_legionella	336	38	46.9	1678	2	Q9P2N9_HUMAN	Q9p2n9_homo sapien
264	38	46.9	255	2	Q54169_STRFR	Q54169_streptomyce	337	38	46.9	2056	2	Q8CHF3_MOUSE	Q8chf3_mus muscucu
265	38	46.9	256	2	Q9S4D5_STRFR	Q9s4d5_streptomyce	338	38	46.9	2058	2	Q5T4S9_HUMAN	Q5t4s9_homo sapien
266	38	46.9	259	2	Q9RRQ0_DEIRA	Q9rrq0_deinococcus	339	38	46.9	2368	2	Q6LFN0_PLAF7	Q6lfn0_plasmodium
267	38	46.9	262	2	Q4WDH7_ASPFU	Q4wdh7_aspergillus	340	38	46.9	2553	2	Q22860_CAEL	Q22860_caenorhabdi
268	38	46.9	267	2	Q6N1W6_RHOPA	Q6n1w6_rhodospheus	341	38	46.9	2553	2	Q51XM7_WAGGR	Q51xm7_magnaporthe
269	38	46.9	285	2	Q63F10_BACCZ	Q63f10_bacillus ce	342	38	46.9	5129	2	Q4S817_TETNG	Q4s817_tetradodon n
270	38	46.9	285	2	Q81U89_BACAN	Q81u89_bacillus an	343	38	46.9	5183	2	Q8TDN5_HUMAN	Q8tdn5_homo sapien
271	38	46.9	295	2	Q8YPB9_ANASP	Q8ypb9_anabaena sp	344	38	46.9	5183	2	Q5T4S7_HUMAN	Q5t4s7_homo sapien
272	38	46.9	310	2	Q64924_MAIZE	Q64924_zea mays (m	345	37.5	46.3	262	2	Q9HT11_PSEAE	Q9ht11_pseudomonas
273	38	46.9	323	2	Q9KFL1_BACHD	Q9kfl1_bacillus ha	346	37.5	46.3	279	2	Q7PM18_ANOGA	Q7pm18_anopheles g
274	38	46.9	337	2	Q41975_MHV68	Q41975_murid herpe	347	37.5	46.3	327	2	Q8TFW5_ASPFU	Q8tfw5_aspergillus
275	38	46.9	342	2	Q4KL87_XENLA	Q4kl87_xenopus lae	348	37.5	46.3	327	2	Q4WSB6_ASPFU	Q4wsb6_aspergillus
276	38	46.9	379	2	Q83DG6_COXBU	Q83dg6_coxiella bu	349	37.5	46.3	364	1	MURG_THETN	Q8r996_thermoanaer
277	38	46.9	381	2	Q5F812_NEIG1	Q5f812_neisseria g	350	37.5	46.3	385	2	Q589Y1_TOBAC	Q589y1_nicotiana t
278	38	46.9	388	2	Q7ZY44_XENLA	Q7zy44_xenopus lae	351	37.5	46.3	485	2	Q4WJ59_ASPFU	Q4wj59_aspergillus
279	38	46.9	401	2	Q971N9_SULTO	Q971n9_sulfolobus l	352	37.5	46.3	584	2	KUPL_LAGJO	Q74ln3_lactobacilli
280	38	46.9	421	1	Y118_TREPA	Q8kj51_rhizobium l	353	37.5	46.3	683	1	Q28483_MACPA	Q28483_macaca fasc
281	38	46.9	422	2	Q5KEZ8_CRYNE	Q5kez8_cryptococcu	354	37.5	46.3	756	2	Q4SR20_TETNG	Q4sr20_tetradodon n
282	38	46.9	424	2	Q98AU7_RHILO	Q98au7_rhizobium l	355	37.5	46.3	786	2	Q4UBD0_THEAN	Q4ubd0_theileria a
283	38	46.9	425	1	G6PD_HELPY	Q58155_treponema p	356	37.5	46.3	829	2	ADA22_XENLA	Q42596_xenopus lae
284	38	46.9	440	2	Q7NJD4_GLOVI	Q98au7_rhizobium l	357	37.5	46.3	935	1	Q51JA3_9VIRU	Q51ja3_apricot pee
285	38	46.9	447	1	DHE4_SALTI	P56110_helicobacte	358	37	45.7	1892	2	Q5C720_SCHJA	Q5c720_schistosoma
286	38	46.9	447	1	DHE4_SALTY	Q7rjd4_gloeobacte	359	37	45.7	49	2	Q7MIR3_VIBYU	Q7mir3_vibrio vuln
287	38	46.9	447	2	Q57P4_SALCH	Q8z6f6_salmonella	360	37	45.7	68	2	Q9YBL4_AERPE	Q9ybl4_aeropyrum p
288	38	46.9	447	2	Q5PHC4_SALPA	P15111_salmonella	361	37	45.7	101	2	Q9YBL4_AERPE	Q9ybl4_aeropyrum p
289	38	46.9	455	2	Q9JN57_STRFR	Q57py4_salmonella	362	37	45.7	108	2	Q5H280_XANOR	Q5h280_xanthomonas
290	38	46.9	460	2	Q5SNC9_CRYNE	Q5phc4_salmonella	363	37	45.7	122	2	Q9QJ75_9RHAB	Q9qjt5_snakehead r
291	38	46.9	460	2	Q5KBR6_CRYNE	Q5jnc9_cryptococcu	364	37	45.7	151	2	Q58EA0_XENLA	Q58ea0_xenopus lae
292	38	46.9	460	2	Q5KBR5_CRYNE	Q5kbr6_cryptococcu	365	37	45.7	167	2	Q90XJ2_9ACTI	Q90xj2_acipenser s
293	38	46.9	465	2	Q8L4X7_ORYSA	Q5kbr5_cryptococcu	366	37	45.7	173	2	Q7VVG1_BORPE	Q7vvg1_bordetella
294	38	46.9	469	2	Q4KCQ9_PSEF5	Q8l4x7_oryza sativ	367	37	45.7	173	2	Q7WBR3_BORBR	Q7wbr3_bordetella
295	38	46.9	469	2	Q89H21_PSEPK	Q4kcq9_pseudomonas	368	37	45.7	177	2	Q67705_9VIRU	Q67705_grapevine v
296	38	46.9	471	2	Q9HWP9_PSEAE	Q8hwp9_pseudomonas	369	37	45.7	188	2	Q6DFJ3_XENLA	Q6dfj3_xenopus lae
297	38	46.9	476	2	Q9REU2_STRFR	Q8zr7_pyrobaculum	370	37	45.7	192	2	Q6SGI7_9BACT	Q6sgi7_uncultured
298	38	46.9	476	2	Q8F8B3_LEPIN	Q9reuz2_streptomyce	371	37	45.7	192	2	Q6SHH7_9BACT	Q6shh7_uncultured
299	38	46.9	480	2	Q61PF3_CAEBR	Q8f8b3_leptospira	372	37	45.7	195	2	Q7ZTE9_PAUPA	Q7zte9_pauxi pauxi
300	38	46.9	505	2	Q6WKZ7_MOUSE	Q6wkz7_mus muscucu	373	37	45.7	195	2	Q7ZTF0_9GALL	Q7ztf0_nothocrax u
301	38	46.9	506	2	Q7PFSK5_MOUSE	Q7pfsk5_mus muscucu	374	37	45.7	195	2	Q7ZTF1_9GALL	Q7ztf1_mitu tubero
302	38	46.9	513	2	Q6UP84_ALCEU	Q6up84_alcaligenes	375	37	45.7	195	2	Q7ZTF2_CRABL	Q7ztf2_crax blumen
303	38	46.9	514	2	Q6UP84_ALCEU	Q6up84_alcaligenes	376	37	45.7	195	2	Q7ZTF4_OREDE	Q7ztf4_oreophasis
304	38	46.9	514	2	Q4J8W0_SULAC	Q4j8w0_sulfolobus	377	37	45.7	195	2	Q7ZTF5_PIPJA	Q7ztf5_pipile jacu
305	38	46.9	516	2	Q89VQ1_BRAJA	Q89vq1_bradyrhizob	378	37	45.7	195	2	Q7ZTF6_9GALL	Q7ztf6_penelopina
306	38	46.9	517	1	VNNL2_DROME	Q8irri_drosophila	379	37	45.7	195	2	Q7ZTF7_9GALL	Q7ztf7_penelope ob
307	38	46.9	519	2	Q84FW2_METEX	Q84fw2_methylobact	380	37	45.7	195	2	Q7ZTF8_9GALL	Q7ztf8_chamaepetes
308	38	46.9	526	2	Q7WMR8_BORBR	Q7wmr8_bordetella	381	37	45.7	196	2	Q7ZTF9_9GALL	Q7ztf9_aburria abu
309	38	46.9	532	2	Q5SR72_CRYNE	Q5sr72_cryptococcu	382	37	45.7	196	2	Q58X02_ORENI	Q58x02_oreochromis
310	38	46.9	543	2	Q65114_ASF	Q65114_african swi	383	37	45.7	198	2	Q5BP11_ARATH	Q5bp11_arabidopsis
311	38	46.9	582	2	Q6C659_YARLI	Q6c659_yarrowia li	384	37	45.7	200	2	Q90XW3_SCOUM	Q90xw3_scopus umbr
312	38	46.9	582	2	Q522A2_CARAU	Q522a2_carassius a	385	37	45.7	202	1	END3_BUCAP	END3_BUCAP
313	38	46.9	582	2	Q7NGP4_GLOVI	Q7ngp4_gloeobacte	386	37	45.7	202	1	END3_BUCAP	END3_BUCAP
314	38	46.9	616	2	Q7NGP4_GLOVI	Q7ngp4_gloeobacte	387	37	45.7	210	1	END3_BUCAP	END3_BUCAP
315	38	46.9	626	2	Q8DKA9_SYNEL	Q8dka9_synechococ	388	37	45.7	211	2	Q6H8Q3_CANFA	Q6h8q3_canis famil
316	38	46.9	648	2	Q8YMH3_ANASP	Q8ymh3_anabaena sp	389	37	45.7	214	2	Q6RYS3_DEIRA	Q6rys3_deinococcus
317	38	46.9	670	2	Q8RC15_PONPY	Q8rc15_pongopygma	390	37	45.7	218	2	Q6ZW87_HUMAN	Q6zw87_homo sapien
318	38	46.9	679	1	Y005_SYNY3	Q5c5560_synechocyc	391	37	45.7	220	2	Q4S895_TETNG	Q4s895_tetradodon n
319	38	46.9	681	1	Y005_SYNY3	Q5c5560_synechocyc	392	37	45.7	221	2	Q801N6_XENLA	Q801n6_xenopus lae
320	38	46.9	683	2	Q6MF18_PARUM	Q6mf18_parachlamyd	393	37	45.7	223	2	Q51EAG_ENTHI	Q51eag_entamoeba h
321	38	46.9	718	2	Q5CUJ3_CRYHO	Q5cuj3_cryptospori	394	37	45.7	227	2	Q65QMB_MANSM	Q65qmb_mannheimia
322	38	46.9	719	2	Q7YYQ4_CRYPV	Q7yyq4_cryptospori	395	37	45.7	231	2	Q8SQF1_NATST	Q8sqf1_natalus str
323	38	46.9	729	2	Q8LMU8_ORYSA	Q8lmu8_oryza sativ	396	37	45.7	238	2	Q8SQF1_NATST	Q8sqf1_natalus str

397	37	45.7	255	2	Q9A2U9 CAUCR	Q9A2U9 caulobacter	470	37	45.7	541	1	CGT_MOUSE	Q64676 mus musculus
398	37	45.7	258	2	Q88QC0_PSEPK	Q88qc0 pseudomonas	471	37	45.7	541	1	CGT_RAT	Q09426 rattus norv
399	37	45.7	266	2	Q5XGL8_XENLA	Q5xgl8 xenopus lae	472	37	45.7	541	2	Q91W57_MOUSE	Q01w57 mus musculus
400	37	45.7	267	2	Q7MYV2_PHOLL	Q7myv2 photorhabd	473	37	45.7	541	2	Q98TB5_CHICK	Q98tb5 gallus gall
401	37	45.7	272	1	PANB_XYLPT	Q9pgr9 xyliella fas	474	37	45.7	543	2	Q80L07_BRARE	Q80lu7 brachydanio
402	37	45.7	272	1	PANB_XYLPT	Q87ew0 xyliella fas	475	37	45.7	569	2	Q6BG99_PARTE	Q6bg99 paramecium
403	37	45.7	272	2	Q6JQO4_BOMOR	Q6lqq4 bombina ori	476	37	45.7	573	2	Q5RK13_RAT	Q5rk13 rattus norv
404	37	45.7	291	2	Q8NIC2_HUMAN	Q8nlc2 homo sapien	477	37	45.7	586	2	Q4SU67_TETNG	Q4su67 tetraodon n
405	37	45.7	291	2	Q86S85_HUMAN	Q86ses homo sapien	478	37	45.7	593	2	Q4MAG8_ASPTU	Q4mag8 aspergillus
406	37	45.7	293	2	Q8BTGL_MOUSE	Q8btgl mus musculus	479	37	45.7	595	2	Q4IU4_GIBZE	Q4iu4 gibberella
407	37	45.7	293	2	Q8BTFL_MOUSE	Q8btfl mus musculus	480	37	45.7	595	2	Q7PHL5_ANOGA	Q7phl5 anopheles g
408	37	45.7	302	2	Q5LQ50_SILPO	Q5lq50 silicibacte	481	37	45.7	602	1	Q4S9P4_TETNG	Q4s9p4 tetraodon n
409	37	45.7	305	2	Q5M905_XENTR	Q5m905 xenopus tro	482	37	45.7	610	1	LYAM2_HORSE	Q95lgl equus cabal
410	37	45.7	308	2	Q8M77_PSEPK	Q8m77 pseudomonas	483	37	45.7	629	2	Q7QC3_ANOGA	Q7qc3 anopheles g
411	37	45.7	310	2	Q38AA7_RHIL0	Q38aa7 rhizobium l	484	37	45.7	638	2	Q4IU8_AZOV1	Q4iu8 azotobacter
412	37	45.7	314	2	Q2I849_CABEL	Q2i849 caenorhabd1	485	37	45.7	645	2	Q6K213_ORISA	Q6k213 oryza sativ
413	37	45.7	318	2	Q6CLH9_KULLA	Q6clh9 kluyveromyc	486	37	45.7	670	2	Q4RL61_TETNG	Q4rl61 tetraodon n
414	37	45.7	318	2	Q4SH90_TETNG	Q4sh80 tetraodon n	487	37	45.7	672	2	Q8A3Q5_BACTN	Q8a3q5 bacteroides
415	37	45.7	321	2	Q4L913_STAHL	Q4l913 staphylococ	488	37	45.7	677	2	Q4RLZ5_TETNG	Q4rlz5 tetraodon n
416	37	45.7	328	2	Q4HLW7_CAMLA	Q4hlw7 campylobact	489	37	45.7	680	2	Q4Z175_BRARE	Q4z175 brachydanio
417	37	45.7	329	2	Q6TBR4_TOXGO	Q6tbr4 toxoplasma	490	37	45.7	688	2	Q6PBB2_BRARE	Q6pb22 brachydanio
418	37	45.7	332	2	Q86UT8_HUMAN	Q86ut8 homo sapien	491	37	45.7	688	2	Q4SH09_TETNG	Q4sh09 tetraodon n
419	37	45.7	333	2	Q7ZWH8_BRARE	Q7zwh8 brachydanio	492	37	45.7	702	2	Q4SH09_TETNG	Q4sh09 tetraodon n
420	37	45.7	333	2	Q4S564_TETNG	Q4s564 tetraodon n	493	37	45.7	716	1	YDOA_SCHPO	Q13730 schizosacch
421	37	45.7	333	2	Q98T53_BRARE	Q98ts3 brachydanio	494	37	45.7	730	2	Q5RFV8_BRARE	Q5rfv8 brachydanio
422	37	45.7	334	2	Q4UI01_THEAN	Q4ui01 theileria a	495	37	45.7	730	2	Q5BKX4_BRARE	Q5bkx4 brachydanio
423	37	45.7	334	2	Q4N7L0_THEPA	Q4n7l0 theileria p	496	37	45.7	749	1	APBA2_PONPY	Q5rd33 pongo pygma
424	37	45.7	338	2	Q6H9N9_YARLI	Q6h9n9 yarrowia li	497	37	45.7	750	1	APBA2_RAT	Q35431 rattus norv
425	37	45.7	350	2	Q8EW16_MYCPE	Q8ew16 mycoplasma	498	37	45.7	776	1	CHS1_CANAL	Q23316 candida alb
426	37	45.7	358	2	Q5FP01_GLUOX	Q5fp01 gluconobact	499	37	45.7	833	1	CWH41_YEAST	P33008 saccharomyc
427	37	45.7	370	2	Q5NG03_FRATT	Q5ng03 francisella	500	37	45.7	835	2	Q4I3L3_GIBZE	Q4i3l3 gibberella
428	37	45.7	372	2	Q6Z786_ORISA	Q6z786 oryza sativ	501	37	45.7	856	2	Q6C0S9_YARLI	Q6c0s9 yarrowia li
429	37	45.7	376	2	Q5ER65_MONAL	Q5er65 monopterus	502	37	45.7	881	2	Q8JAG5_SIVCZ	Q8jag5 chimpanzee
430	37	45.7	387	1	ILF2_BRARE	Q6nz06 brachydanio	503	37	45.7	882	2	Q6N313_RHOPA	Q6n313 rhodopseudo
431	37	45.7	408	2	Q53J99_ORISA	Q53j99 oryza sativ	504	37	45.7	931	2	Q5E319_VIBF1	Q5e319 vibrio fisc
432	37	45.7	413	2	Q7YRC5_BOVIN	Q7yrc5 bos taurus	505	37	45.7	938	2	Q4QI88_LEIMA	Q4ql88 leishmania
433	37	45.7	414	2	Q5N1M6_SVNP6	Q5nlm6 synechococc	506	37	45.7	946	2	Q57V94_9TRYP	Q57v94 trypanosoma
434	37	45.7	415	2	Q6I0N9_XENLA	Q6inn9 xenopus lae	507	37	45.7	980	2	Q5U167_DROME	Q5u167 drosophila
435	37	45.7	416	2	Q9XG64_CABEL	Q9xx64 caenorhabd1	508	37	45.7	1005	2	Q5K995_CRYNE	Q5k995 cryptococcu
436	37	45.7	424	2	Q5N1N0_FRATT	Q5nn10 francisella	509	37	45.7	1005	2	Q5SKA7_CRYNE	Q5ska7 cryptococcu
437	37	45.7	434	2	Q4RY62_TETNG	Q4ry62 tetraodon n	510	37	45.7	1012	2	Q6BZ46_DEBHA	Q6bz46 debaryomyc
438	37	45.7	436	2	Q7UV93_RHOBA	Q7uv93 rhodopirell	511	37	45.7	1022	2	Q6CS24_KULLA	Q6cs24 kluyveromyc
439	37	45.7	438	2	Q6SH08_9BACT	Q6sh08 uncultured	512	37	45.7	1026	2	Q5A594_CANAL	Q5a594 candida alb
440	37	45.7	444	2	Q5LNM2_MAGGR	Q5lnm2 magnaporthe	513	37	45.7	1026	2	Q9URM1_CANAL	Q9urml candida alb
441	37	45.7	447	2	Q3P0F8_AZOSE	Q3p0f8 azarcus sp	514	37	45.7	1044	2	Q8QUS6_9VIRU	Q8qus6 infectious
442	37	45.7	454	2	Q864N3_NATST	Q864n3 natalus str	515	37	45.7	1045	2	Q8BM83_MOUSE	Q8bm83 m mus muscu
443	37	45.7	454	2	Q53A16_9CHIR	Q53a16 natalus tum	516	37	45.7	1052	2	Q4KSC1_9VIRU	Q4ksc1 orange-spot
444	37	45.7	454	2	Q53A17_9CHIR	Q53a17 natalus sat	517	37	45.7	1053	2	Q5YF54_9VIRU	Q5yf54 rock bream
445	37	45.7	454	2	Q53A18_9CHIR	Q53a18 natalus jam	518	37	45.7	1055	2	Q6FKY0_CANGA	Q6fky0 candida gla
446	37	45.7	454	2	Q53A19_9CHIR	Q53a19 natalus maj	519	37	45.7	1105	2	Q55PS2_DICDI	Q55fs2 dictyosteli
447	37	45.7	454	2	Q53A20_9CHIR	Q53a20 chilonatalu	520	37	45.7	1106	2	Q4Z291_CHICK	Q4z291 gallus gall
448	37	45.7	454	2	Q53A21_9CHIR	Q53a21 nyctetellus	521	37	45.7	1142	2	Q26615_STRPU	Q26615 strongyloce
449	37	45.7	455	2	Q7VZQ0_BORPE	Q7vzq0 bordetella	522	37	45.7	1168	2	Q60XC0_CAEBR	Q60xc0 caenorhabd1
450	37	45.7	455	2	Q7W5B2_BORPA	Q7w5b2 bordetella	523	37	45.7	1212	2	Q4RLZ8_TETNG	Q4rlz8 tetraodon n
451	37	45.7	455	2	Q7WCU4_BORBR	Q7wcu4 bordetella	524	37	45.7	1244	1	MXI1_CABEL	Q95991 caenorhabd1
452	37	45.7	461	2	Q66K22_DICDI	Q66k22 dictyosteli	525	37	45.7	1279	2	Q86PE6_DROME	Q86pe6 drosophila
453	37	45.7	468	2	Q52KU4_XENLA	Q52ku4 xenopus lae	526	37	45.7	1296	2	Q61ZQ5_CAEBR	Q61zq5 caenorhabd1
454	37	45.7	469	1	GLNA_PROVU	P28786 proteus vul	527	37	45.7	1307	2	Q9GT78_PLAFA	Q9gt78 plasmodium
455	37	45.7	473	2	Q9LJ71_ARATH	Q9lj71 arabidopsis	528	37	45.7	1322	2	Q7X7E9_ORISA	Q7x7e9 oryza sativ
456	37	45.7	473	2	Q7WMB7_BORPA	Q7wmb7 bordetella	529	37	45.7	1368	2	Q8I199_PLAET	Q8i199 plasmodium
457	37	45.7	473	2	Q7WMB9_BORBR	Q7wmb9 bordetella	530	37	45.7	1411	2	Q5ARI1_BWENI	Q5ari1 aspergillus
458	37	45.7	478	2	Q74H74_GROSL	Q74h74 geobacter s	531	37	45.7	1460	2	Q54HC6_DICDI	Q54hc6 dictyosteli
459	37	45.7	489	2	Q569D6_XENTR	Q569d6 xenopus tro	532	37	45.7	1479	2	Q54BFD0_BWENI	Q5bfd0 aspergillus
460	37	45.7	490	2	Q8RZ19_ORISA	Q8rz19 oryza sativ	533	37	45.7	1544	2	Q3W2F2_DROME	Q3w2f2 drosophila
461	37	45.7	491	2	Q62F12_BURMA	Q62f12 burkholderi	534	37	45.7	1664	2	Q9TVQ2_CABEL	Q9tvq2 caenorhabd1
462	37	45.7	491	2	Q63VC1_BURPS	Q63vc1 burkholderi	535	37	45.7	1912	1	CHD4_HUMAN	Q14839 homo sapien
463	37	45.7	493	2	Q65GCO_BACLD	Q65gc0 bacillus li	536	37	45.7	1915	1	CHD4_MOUSE	Q6pdp2 mus musculu
464	37	45.7	496	2	Q6C7M6_YARLI	Q6c7m6 yarrowia li	537	37	45.7	1945	2	Q5DTF7_MOUSE	Q5dtp7 mus musculu
465	37	45.7	498	2	Q9BDZ8_BOVIN	Q9bdz8 bos taurus	538	37	45.7	2248	2	Q4S447_TETNG	Q4s447 tetraodon n
466	37	45.7	505	2	Q4SSLO_TETNG	Q4ssl0 tetraodon n	539	37	45.7	2386	1	RAD3_SCHPO	Q20299 schizosacch
467	37	45.7	518	2	Q6NBU5_RHOPA	Q6nbu5 rhodopseudo	540	37	45.7	2529	2	Q7QJP5_ANOGA	Q7qjp5 anopheles g
468	37	45.7	530	2	Q8NSQ1_CORGL	Q8nsq1 corynebacte	541	37	45.7	2531	1	NOTC1_RAT	Q70708 rattus norv
469	37	45.7	540	2	Q9E201_9ALPH	Q9e201 cercopithe	542	37	45.7	2820	1	NP1_RAT	P97526 rattus norv

543	37	45.7	2820	2	Q5SY11_MOUSE	Q5ey11 mus musculus	616	36	44.4	271	1	KKAB_ECOLI	P14509 escherichia
544	37	45.7	2839	1	NF1_HUMAN	P21359 homo sapien	617	36	44.4	271	2	Q75IR0_ORYSA	Q75ir0 oryza sativ
545	37	45.7	2841	1	NF1_MOUSE	Q04690 mus musculu	618	36	44.4	274	2	Q6UK87_9CAUD	Q6uk87 burkholderi
546	37	45.7	2841	1	Q5SY12_MOUSE	Q5ey12 mus musculu	619	36	44.4	275	2	Q74SW8_YERPE	Q74sw8 yersinia pe
547	37	45.7	4751	2	Q4QID5_LEIMA	Q4qid5 leishmania	620	36	44.4	282	1	ESTD_HUMAN	P10768 homo sapien
548	37	45.7	7047	2	Q5BA85_EMENI	Q5ba85 aspergillus	621	36	44.4	282	1	ESTD_MOUSE	Q5tbr0 mus musculu
549	36.5	45.1	255	2	Q5W960_BOTJA	Q5w960 bothrops ja	622	36	44.4	282	2	Q5TBV0_HUMAN	Q5tbr0 homo sapien
550	36.5	45.1	262	2	Q4J5G1_AZQVI	Q4j5g1 azotobacter	623	36	44.4	282	2	Q8ZD66_YERPE	Q8zdb6 yersinia pe
551	36.5	45.1	350	1	RA51B_MOUSE	Q5y19 mus musculu	624	36	44.4	282	2	Q668T0_XENLA	Q668t0 xenopus lae
552	36.5	45.1	379	2	Q9M1E0_ARATH	Q9mle0 arabidopsis	625	36	44.4	286	2	Q8AXC0_XENLA	Q8axc0 xenopus lae
553	36.5	45.1	600	2	Q4FY16_LEIMA	Q4fy16 leishmania	626	36	44.4	287	2	Q9ZGJ0_PHRCA	Q9zgj0 ehrlichia c
554	36.5	45.1	694	2	Q9LKR4_ARATH	Q9lkr4 arabidopsis	627	36	44.4	293	2	Q4S9G9_TETNG	Q4s9g9 tetraodon n
555	36.5	45.1	1461	2	Q4HUT6_GIBZE	Q4hut6 gibberella	628	36	44.4	294	2	Q5CUR6_CRYPV	Q5cur6 cryptospori
556	36.5	45.1	7599	2	Q4IKK3_GIBZE	Q4ike3 gibberella	629	36	44.4	298	2	Q8TUX5_METKA	Q8tux5 methanopyru
557	36	44.4	63	2	Q8GQ95_PSEAE	Q8gg95 pseudomonas	630	36	44.4	298	2	Q6K1S1_ORYSA	Q6k1s1 oryza sativ
558	36	44.4	68	2	Q7UVN1_RHOBA	Q7uvn1 rhodopirell	631	36	44.4	302	2	Q8EEB4_SHEON	Q8eeeb4 shewanella
559	36	44.4	88	2	Q9XGV3_GIAMI	Q9xgv3 orobancha a	632	36	44.4	302	2	Q8BGE4_MOUSE	Q8bge4 m hypotheti
560	36	44.4	91	2	Q8EBT0_SHEON	Q8ebt0 shewanella	633	36	44.4	303	1	K0247_HUMAN	Q92537 homo sapien
561	36	44.4	98	2	Q4XR92_PLACH	Q4xr92 plasmodium	634	36	44.4	309	2	Q9D9N6_MOUSE	Q9d9n6 mus musculu
562	36	44.4	101	2	Q854B4_9CAUD	Q854b4 mycobacteri	635	36	44.4	316	2	Q5NNR0_ZYMMO	Q5nnr0 zymomonas m
563	36	44.4	106	2	Q4SNK1_TETNG	Q4snk1 tetraodon n	636	36	44.4	318	2	Q5AEU1_CANAL	Q5aeul candida alb
564	36	44.4	118	2	P97731_9MORI	P97731 rattus sp.	637	36	44.4	320	2	Q5BPL5_ARATH	Q5bpl5 arabidopsis
565	36	44.4	124	2	Q55PE4_CRYNE	Q55pe4 cryptococcu	638	36	44.4	320	2	Q5XV84_ARATH	Q5xv84 arabidopsis
566	36	44.4	132	2	Q8D9U4_VIBVU	Q8d9u4 vibrio vuln	639	36	44.4	321	2	Q4LNE9_BURK	Q4ln69 burkholderi
567	36	44.4	148	2	Q651P1_ORYSA	Q651p1 oryza sativ	640	36	44.4	321	2	Q8COM5_MOUSE	Q8com5 mus musculu
568	36	44.4	156	2	Q8D9U4_VIBVU	Q8d9u4 vibrio vuln	641	36	44.4	324	2	Q6ZQF4_MOUSE	Q6zqf4 mus musculu
569	36	44.4	165	1	Y868_PICTO	Q8qnu5 synecococc	642	36	44.4	324	2	Q6ZQF4_MOUSE	Q6zqf4 mus musculu
570	36	44.4	167	2	Q5DBI0_SCHJA	Q5dbi0 schistosoma	643	36	44.4	332	1	BNIP1_MOUSE	Q924v5 cavpo
571	36	44.4	167	2	Q8XMA2_CLOPE	Q8xma2 clostridium	644	36	44.4	334	1	ADD1_VIBPA	Q924v5 cavpo
572	36	44.4	171	2	Q4Q417_9ASPA	Q4q417 narciessus h	645	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
573	36	44.4	174	2	Q4Q417_9ASPA	Q4q417 narciessus h	646	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
574	36	44.4	175	2	Q5WB18_BACSK	Q5wb18 bacillus cl	647	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
575	36	44.4	177	2	Q7T465_9VIRU	Q7t465 grapevine v	648	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
576	36	44.4	177	2	Q8B3K5_9VIRU	Q8b3k5 grapevine v	649	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
577	36	44.4	188	2	Q5VQZ7_ORYSA	Q5vqz7 oryza sativ	650	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
578	36	44.4	189	2	Q7PCP3_ANOGA	Q7pcp3 anopheles g	651	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
579	36	44.4	193	2	Q7Q710_ANOGA	Q7q710 anopheles g	652	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
580	36	44.4	194	2	Q7ZTGO_MEGRE	Q7ztgo megapodius	653	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
581	36	44.4	196	2	Q5VIH9_PHACC	Q5vih9 phasianus c	654	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
582	36	44.4	199	2	Q6VSP5_9PASS	Q6vsp5 toxosoma l	655	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
583	36	44.4	200	2	P87425_9PASS	P87425 acanthisitt	656	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
584	36	44.4	202	2	Q8QH51_9PASS	Q8qh51 mimus patag	657	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
585	36	44.4	202	2	Q8QH51_9PASS	Q8qh51 mimus patag	658	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
586	36	44.4	204	2	Q8KJAC_RHILO	Q8kjac rhizobium l	659	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
587	36	44.4	221	2	Q8AXC1_XENLA	Q8axc1 xenopus lae	660	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
588	36	44.4	222	2	Q7Q710_ANOGA	Q7q710 anopheles g	661	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
589	36	44.4	222	2	Q6CZV7_ERWCT	Q6czv7 erwinia car	662	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
590	36	44.4	223	2	Q7W711_BORPA	Q7w711 bordetella	663	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
591	36	44.4	223	2	Q7W711_BORPA	Q7w711 bordetella	664	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
592	36	44.4	223	2	Q9S3M3_BORPE	Q9s3m3 bordetella	665	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
593	36	44.4	223	2	Q5XHS5_CRYNE	Q5xhs5 cryptococcu	666	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
594	36	44.4	229	2	Q5XMF9_CRYNE	Q5xmf9 cryptococcu	667	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
595	36	44.4	234	2	Q4JAE7_AZQVI	Q4jae7 azotobacter	668	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
596	36	44.4	238	1	FLGH_BUCAI	P57425 buchnera ap	669	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
597	36	44.4	242	2	Q7PSR1_ANOGA	Q7psr1 anopheles g	670	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
598	36	44.4	244	2	Q73CS3_BACCL	Q73cs3 bacillus ce	671	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
599	36	44.4	244	2	Q7WRD3_BORBR	Q7wrd3 bordetella	672	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
600	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	673	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
601	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	674	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
602	36	44.4	244	2	Q7WRD3_BORBR	Q7wrd3 bordetella	675	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
603	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	676	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
604	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	677	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
605	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	678	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
606	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	679	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
607	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	680	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
608	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	681	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
609	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	682	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
610	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	683	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
611	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	684	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
612	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	685	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
613	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	686	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
614	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	687	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo
615	36	44.4	244	2	Q7W2GA_BORPA	Q7w2ga bordetella	688	36	44.4	335	1	IL12B_MOUSE	Q924v5 cavpo

689	36	44.4	459	2	Q4U3V5_CRYPA	Q4u3v5_cryphonectr	762	36	44.4	651	2	Q712D5_ORYSA	Q712d5_oryza sativ
690	36	44.4	460	2	Q9AU2_ORYSA	Q9auv2_oryza sativ	763	36	44.4	651	2	Q852P4_ORYSA	Q852p4_oryza sativ
691	36	44.4	461	2	Q9AU9_ORYSA	Q9auu9_oryza sativ	764	36	44.4	654	2	Q8BV4_MOUSE	Q8bv4_mus musculu
692	36	44.4	462	2	Q6FB37_ACTIAD	Q6fb37_actinetobact	765	36	44.4	661	2	Q512S1_ENTHI	Q512s1_entameoba h
693	36	44.4	463	2	Q67B63_9CHLR	Q67b63_dehalococco	766	36	44.4	662	2	Q783U9_NEUCR	Q783u9_neurospora
694	36	44.4	482	2	Q67B92_9CHLR	Q67b92_dehalococco	767	36	44.4	664	2	Q9GRV2_CABEL	Q9grv2_caenorhabdi
695	36	44.4	481	2	Q67B92_9CHLR	Q67b92_dehalococco	768	36	44.4	667	2	Q4NBX4_9MICC	Q4nbx4_arthrobacte
696	36	44.4	490	2	Q7V413_PROMM	Q7v413_prochloroco	769	36	44.4	674	2	Q4HY53_GIBZE	Q4hy53_gibberella
697	36	44.4	497	2	Q6Z1A7_ORYSA	Q6z1a7_oryza sativ	770	36	44.4	679	2	Q9K95_BACHD	Q9k95_bacillus ha
698	36	44.4	497	2	Q9FYU7_BRANA	Q9fyu7_brassica na	771	36	44.4	684	2	Q9K95_BACHD	Q9k95_bacillus ha
699	36	44.4	497	2	Q52TA3_ENTAG	Q52ta3_enterobacte	772	36	44.4	721	2	Q7S7G6_NEUCR	Q7s7g6_neurospora
700	36	44.4	505	2	Q9VZJ4_DROME	Q9vzj4_drosophila	773	36	44.4	745	2	Q4FPP5_9RICK	Q4fpp5_candidatus
701	36	44.4	505	2	Q8BX48_MOUSE	Q8bx48_mus musculu	774	36	44.4	751	2	Q580F4_TRYANOS	Q580f4_trypanosoma
702	36	44.4	510	2	Q4IMA4_GIBZE	Q4ima4_gibberella	775	36	44.4	764	2	Q7R0K3_GIALA	Q7r0k3_giardia lam
703	36	44.4	513	2	Q9D118_MOUSE	Q9d118_mus musculu	776	36	44.4	764	2	P93756_ARATH	P93756_arabidopsais
704	36	44.4	516	2	Q7X871_ORYSA	Q7x871_oryza sativ	777	36	44.4	767	2	P93756_ARATH	P93756_arabidopsais
705	36	44.4	517	2	Q9VJ45_DROME	Q9vj45_drosophila	778	36	44.4	777	2	Q7S1F1_NEUCR	Q7s1f1_neurospora
706	36	44.4	522	1	COR1_SCHPO	O13688_schizosacch	779	36	44.4	780	2	Q4RL36_TETNG	Q4rl36_tetradodon n
707	36	44.4	527	2	Q52F89_MAGGR	Q52f89_magnaporthe	780	36	44.4	786	2	Q9FVW7_ARATH	Q9fvw7_arabidopsais
708	36	44.4	528	2	Q5B6G5_WMENI	Q5b6g5_aspergillus	781	36	44.4	796	2	Q6B228_DEBHA	Q6b228_debaryomyce
709	36	44.4	530	2	O16276_CABEL	O16276_caenorhabdi	782	36	44.4	803	2	Q9CB85_MYCLE	Q9cb85_mycobacteri
710	36	44.4	531	2	Q516K3_ENTHI	Q516k3_entameoba h	783	36	44.4	818	2	Q6CNQ6_KLULA	Q6cng6_kluyveromyc
711	36	44.4	534	2	Q21603_CABEL	Q21603_caenorhabdi	784	36	44.4	820	2	Q8GX69_ARATH	Q8gx69_arabidopsais
712	36	44.4	535	2	O18009_CAEEL	O18009_caenorhabdi	785	36	44.4	820	2	Q4HGH7_CAMCO	Q4hgh7_campylobact
713	36	44.4	535	2	O74BV3_GEOSL	O74bv3_geobacter s	786	36	44.4	821	2	O5CVP1_CRYPV	O5cyp1_cryptospori
714	36	44.4	537	2	Q8DIX3_SYNEL	Q8dix3_synechococc	787	36	44.4	821	2	O5CP41_CRYHO	O5cp41_cryptospori
715	36	44.4	537	2	Q5I916_BHARE	Q5i916_brachydanio	788	36	44.4	821	2	P72351_MYCLE	P72351_mycobacteri
716	36	44.4	538	2	Q4FM39_9RICK	Q4fm39_candidatus	789	36	44.4	826	2	P72351_MYCLE	P72351_mycobacteri
717	36	44.4	541	1	CGT_HUMAN	Q16880_homo sapien	790	36	44.4	826	2	O5YJ18_9HERP	O5yj18_fibropapill
718	36	44.4	542	1	PVRG_LEPIC	Q172846_leptospira	791	36	44.4	828	2	O5YJ19_9HERP	O5yj19_fibropapill
719	36	44.4	542	1	PVRG_LEPIN	Q8f3j3_leptospira	792	36	44.4	828	2	O5YJ17_9HERP	O5yj17_fibropapill
720	36	44.4	542	1	PVRG_RHILO	Q98mf0_rhizobium l	793	36	44.4	828	2	O5YJ20_9HERP	O5yj20_fibropapill
721	36	44.4	542	2	Q7PYJ9_ANOGA	Q7pyj9_anopheles g	794	36	44.4	828	2	O5YJ32_9HERP	O5yj32_fibropapill
722	36	44.4	543	2	Q6P8Y4_USTMA	Q6p8y4_ustilago ma	795	36	44.4	828	2	O5YJ33_9HERP	O5yj33_fibropapill
723	36	44.4	543	2	Q6N5T4_RHOPA	Q6n5t4_rhodopseudo	796	36	44.4	839	2	O5YJ34_9HERP	O5yj34_fibropapill
724	36	44.4	543	2	Q89K05_BRAJA	Q89k05_bradyrhizob	797	36	44.4	852	2	O5YJ43_9HERP	O5yj43_fibropapill
725	36	44.4	548	2	O7VQK5_CANBF	Q7vqk5_candidatus	798	36	44.4	854	2	O5Y967_9HERP	O5y967_fibropapill
726	36	44.4	549	2	Q8RN79_STRAT	Q8rn79_streptomyce	799	36	44.4	864	2	O95Q18_CAEEL	O95q18_caenorhabdi
727	36	44.4	555	2	Q8ZF10_YERPE	Q8zf10_yersinia pe	800	36	44.4	866	2	Q8EF55_SHEON	Q8ef55_shewanella
728	36	44.4	555	2	Q669X4_YERPS	Q669x4_yersinia ps	801	36	44.4	877	2	Q7N458_PHOUL	Q7n458_photorhabdu
729	36	44.4	563	2	Q8NG04_HUMAN	Q8ng04_homo sapien	802	36	44.4	886	2	O8JAH4_CHIMPANZEE	O8jah4_chimpanzee
730	36	44.4	564	2	Q8HR52_STAEO	Q8hr52_staphylococ	803	36	44.4	887	2	O7NHY5_CHROMOBACTE	O7nhy5_chromobacte
731	36	44.4	564	2	Q8CQ28_STAEP	Q8cq28_staphylococ	804	36	44.4	890	1	ADHE_ECO57	ADHE_Escherichia
732	36	44.4	569	2	Q8SYK5_TETNG	Q8syk5_tetradodon n	805	36	44.4	890	1	ADHE_ECOLI	ADHE_Escherichia
733	36	44.4	570	2	Q8AEF5_CANAL	Q8aef5_candida alb	806	36	44.4	891	2	Q7DL18_ECOLI	Q7dl18_escherichia
734	36	44.4	572	1	TMPS7_HUMAN	Q7ryt8_homo sapien	807	36	44.4	891	2	Q8FHA2_ECOL6	Q8fha2_escherichia
735	36	44.4	572	1	DPOL_MOUSE	Q9qxe2_mus musculu	808	36	44.4	891	2	Q6DAR4_ERWCT	Q6dar4_erwinia car
736	36	44.4	573	1	DPOL_MOUSE	Q9qxe2_mus musculu	809	36	44.4	892	2	Q83RN2_SHIFL	Q83rn2_shigella fl
737	36	44.4	577	2	Q967G8_NIPBR	Q967g8_nippostrong	810	36	44.4	897	2	Q95Q17_CAEEL	Q95q17_caenorhabdi
738	36	44.4	578	2	Q4P036_USTMA	Q4p036_ustilago ma	811	36	44.4	928	1	O6VG43_SIVCZ	O6vg43_chimpanzee
739	36	44.4	593	2	Q84H36_9BURK	Q84h36_comamonas s	812	36	44.4	932	1	VGLB_BHV1P	VGLB_bovine herp
740	36	44.4	595	2	Q4LU16_9BURK	Q4lu16_burkholderi	813	36	44.4	938	2	VGLB_BHV1P	VGLB_bovine herp
741	36	44.4	595	2	Q62MM6_BURMA	Q62mm6_burkholderi	814	36	44.4	938	2	O580W5_9TRYP	O580w5_trypanosoma
742	36	44.4	595	2	Q63X89_BURPS	Q63x89_burkholderi	815	36	44.4	938	2	O51BQ2_ENTHI	O51bq2_entameoba h
743	36	44.4	601	1	MBHL_AZCH	P18191_azotobacter	816	36	44.4	938	2	Q722D3_LISMF	Q722d3_listeria mo
744	36	44.4	601	1	MBHL_AZCH	P18191_azotobacter	817	36	44.4	938	2	Q8Y8V7_LISMO	Q8y8v7_listeria mo
745	36	44.4	602	1	MBHL_AZCHI	P18191_azotobacter	818	36	44.4	938	2	Q92DN5_LISIN	Q92dn5_listeria in
746	36	44.4	602	1	MBHL_AZCHI	P18191_azotobacter	819	36	44.4	938	2	Q4IM19_GIBZE	Q4im19_gibberella
747	36	44.4	602	1	Q871Y6_NEUCR	Q871y6_neurospora	820	36	44.4	959	2	Q6Z572_ORYSA	Q6z572_oryza sativ
748	36	44.4	602	2	Q4IU00_AZOVI	Q4iu00_azotobacter	821	36	44.4	1022	2	Q4UHY6_THEAN	Q4uhy6_theileria a
749	36	44.4	618	2	Q55FC8_DICDI	Q55fc8_dicotyosteli	822	36	44.4	1027	2	Q9BW22_HUMAN	Q9bw22_homo sapien
750	36	44.4	619	2	Q8ZMW9_HUMAN	Q8zmw9_homo sapien	823	36	44.4	1027	2	Q86VS3_HUMAN	Q86va3_homo sapien
751	36	44.4	621	2	Q9H9Y1_HUMAN	Q9h9y1_homo sapien	824	36	44.4	1032	2	Q4N7M3_THEPA	Q4n7m3_theileria p
752	36	44.4	623	2	Q4RMH4_TETNG	Q4rmh4_tetradodon n	825	36	44.4	1035	2	Q7YIS2_CAEEL	Q7yis2_caenorhabdi
753	36	44.4	634	2	Q517K1_ENTHI	Q517k1_entameoba h	826	36	44.4	1053	2	Q96KJ4_HUMAN	Q96kj4_homo sapien
754	36	44.4	636	2	Q22040_CABEL	Q22040_caenorhabdi	827	36	44.4	1056	1	K125_ARATH	K125_arabidopsais
755	36	44.4	636	2	Q9CAR2_APATH	Q9car2_arabidopsais	828	36	44.4	1074	2	O45878_CAEEL	O45878_caenorhabdi
756	36	44.4	642	2	Q99P70_MOUSE	Q99p70_mus musculu	829	36	44.4	1093	2	O640N3_MOUSE	O640n3_mus musculu
757	36	44.4	642	2	Q9D0F1_MOUSE	Q9d0f1_mus musculu	830	36	44.4	1102	2	Q55WU4_CRYNE	Q55wu4_cryptococcu
758	36	44.4	646	2	Q4WAQ3_ASPPU	Q4waq3_aspergillus	831	36	44.4	1102	2	O5KJF4_CRYNE	O5kjf4_cryptococcu
759	36	44.4	648	2	Q51LH4_BACFN	Q51lh4_bacteroides	832	36	44.4	1156	2	O61137_LEIDO	O61137_leishmania
760	36	44.4	648	2	Q64Y02_BACFR	Q64y02_bacteroides	833	36	44.4	1157	2	O4Q2M2_LEIMA	O4q2m2_leishmania
761	36	44.4	651	2	Q70YK6_ORYSA	Q70yk6_oryza sativ	834	36	44.4	1161	2	Q8J213_KLULA	Q8j213_kluyveromyc

835	36	44.4	1235	1	RENT1 ARATH	Q9fjr0 arabidopsis	908	35	43.2	108	2	Q50B2_PSESY	Q50B2 pseudomonas
836	36	44.4	1243	2	Q8S3K7 ARATH	Q8sk37 arabidopsis	909	35	43.2	109	2	Q8RPAB_PSEST	Q8rpab pseudomonas
837	36	44.4	1310	2	Q4Q2H5 LEIMA	Q4q2h5 leishmania	910	35	43.2	109	2	Q88B13_PSESM	Q88b13 pseudomonas
838	36	44.4	1318	2	Q6ZUA9 HUMAN	Q6zu9a homo sapien	911	35	43.2	113	2	Q9BIF1_9BILA	Q9bif1 cooperia pu
839	36	44.4	1341	2	Q9UIB5 LEIMA	Q9ules leishmania	912	35	43.2	113	2	Q9BIF2_9BILA	Q9bif2 cooperia pu
840	36	44.4	1387	2	Q611T5 CAEBR	Q611t5 caenorhabdi	913	35	43.2	113	2	Q8BIF3_9BILA	Q8bif3 cooperia pu
841	36	44.4	1392	2	Q6LF61 PLAF7	Q6l1f61 plasmodium	914	35	43.2	123	2	Q88IM8_PSESM	Q88im8 pseudomonas
842	36	44.4	1491	2	Q4ID49 GIBZE	Q4id49 gibberella	915	35	43.2	124	2	Q5BXC4_SCHJA	Q5bxc4 schistosoma
843	36	44.4	1491	2	Q5ATW0_EMENI	Q5att0 aspergillus	916	35	43.2	125	2	Q5MK13_9ASPA	Q5mk13 lycoris sp.
844	36	44.4	2848	2	Q7RY33_NEUCR	Q7ry33 neurospora	917	35	43.2	127	2	Q9TV08_CANFA	Q9tv08 canis fami
845	36	44.4	3570	2	Q7Q737_ANOGA	Q7q737 anopheles g	918	35	43.2	128	2	Q8HY59_RABIT	Q8hy59 oryctolagus
846	36	44.4	4144	2	Q8WN22_CANFA	Q8wn22 canis fami	919	35	43.2	131	2	Q6BNN2_DEBHA	Q6bnn2 debaryomyce
847	35.5	43.8	88	2	Q9UKK0_HUMAN	Q9ukko homo sapien	920	35	43.2	131	2	Q8VTN8_HELPY	Q8vtn8 helicobacte
848	35.5	43.8	119	2	Q5CCV1_9GEMI	Q5ccv1 tomato leaf	921	35	43.2	132	2	Q4HAD6_9DEIO	Q4had6 deinococcus
849	35.5	43.8	123	2	Q5CCV2_9GEMI	Q5ccv2 tomato leaf	922	35	43.2	140	2	Q8VTN4_HELPY	Q8vtn4 helicobacte
850	35.5	43.8	158	2	Q8C9S9_MOUSE	Q8c9s9 mus musculu	923	35	43.2	140	2	Q8VTP4_HELPY	Q8vtp4 helicobacte
851	35.5	43.8	159	2	Q4RV20_TETNG	Q4rv20 tetraodon n	924	35	43.2	149	1	Y2595_SULSO	Q97vm6 sulfolobus
852	35.5	43.8	165	2	Q58GM4_9HIV1	Q58gm4 human immu	925	35	43.2	149	2	Q9VMX7_DROME	Q9vmx7 drosophila
853	35.5	43.8	204	2	Q54H08_DICDI	Q54h08 dictyosteli	926	35	43.2	149	2	Q4RSP6_TETNG	Q4rsp6 tetraodon n
854	35.5	43.8	257	2	Q75O16_9GEMI	Q75o16 tomato leaf	927	35	43.2	150	2	Q9ZY20_9MOLLI	Q9zy20 roessia palp
855	35.5	43.8	257	2	Q80S73_9GEMI	Q80s73 pseudomonas	928	35	43.2	160	2	Q7U6B3_SYNXP	Q7u6b3 synechococc
856	35.5	43.8	263	2	Q4ZLL5_PSESM	Q4zll5 pseudomonas	929	35	43.2	160	2	Q5N479_SYNXP	Q5n479 synechococc
857	35.5	43.8	263	2	Q87T85_PSESM	Q87t85 pseudomonas	930	35	43.2	162	2	Q5MK12_9ASPA	Q5mk12 lycoris sp.
858	35.5	43.8	264	2	Q8N3V6_HUMAN	Q8n3v6 homo sapien	931	35	43.2	166	2	Q4QA21_LEIMA	Q4qa21 leishmania
859	35.5	43.8	265	2	Q898C6_CLOTE	Q898c6 clostridium	932	35	43.2	166	2	Q9N9V6_LEIMA	Q9n9v6 leishmania
860	35.5	43.8	265	2	Q8C976_MOUSE	Q8c976 mus musculu	933	35	43.2	166	2	Q8VLE2_HELPY	Q8vle2 helicobacte
861	35.5	43.8	273	2	Q4RFL4_TETNG	Q4rfl4 tetraodon n	934	35	43.2	167	2	Q5VMQ3_HUMAN	Q5vmq3 homo sapien
862	35.5	43.8	311	2	Q9VLM8_DROME	Q9vlm8 drosophila	935	35	43.2	167	2	Q4HJL7_CAMLA	Q4hjl7 campylobact
863	35.5	43.8	334	1	GBLP_ORYSA	P49027 oryza sativ	936	35	43.2	170	2	Q50VB7_ENTHI	Q50vb7 entamoeba h
864	35.5	43.8	339	2	Q5AFI5_CANAL	Q5afi5 candida alb	937	35	43.2	172	2	Q51WS9_MAGGR	Q51ws9 magnaporthe
865	35.5	43.8	342	2	Q5XLY0_GINBI	Q5xly0 ginkgo bilo	938	35	43.2	177	2	Q4HT57_TETNG	Q4ht57 campylobact
866	35.5	43.8	360	2	Q5SMK6_ORYSA	Q5smk6 oryza sativ	939	35	43.2	184	2	Q4T246_CAMP	Q4t246 tetraodon n
867	35.5	43.8	363	1	MURG_LISIN	Q929y2 listeria in	940	35	43.2	185	2	Q7P056_CHRVO	Q7p056 chromobacte
868	35.5	43.8	650	1	KUP1_LACAC	Q8lms3 drosophila	941	35	43.2	186	1	KAD_TROWB	Q83160 tropheryma
869	35.5	43.8	682	2	Q9STF4_ARATH	Q9stf4 arabidopsis	942	35	43.2	186	1	KAD_TROWT	Q83g05 tropheryma
870	35.5	43.8	682	2	Q9STF4_ARATH	Q9stf4 arabidopsis	943	35	43.2	186	1	KAD_TROWT	Q83g05 tropheryma
871	35.5	43.8	697	2	Q4R6Q6_MACFA	Q4r6q6 macaca fasc	944	35	43.2	189	2	Q5VSP7_9PASS	Q5vsp7 thannolea c
872	35.5	43.8	624	2	Q8SZ15_DROME	Q8sz15 drosophila	945	35	43.2	189	2	Q65Q08_MANSM	Q65qq8 mannhelmia
873	35.5	43.8	650	1	KUP1_LACAC	Q8lms3 drosophila	946	35	43.2	191	2	Q5SDB2_VIBAN	Q5sdb2 vibrio angu
874	35.5	43.8	682	2	Q9STF4_ARATH	Q9stf4 arabidopsis	947	35	43.2	192	2	Q731S9_WOLPM	Q731s9 wolbachia p
875	35.5	43.8	906	1	ADA22_HUMAN	Q9p0k1 homo sapien	948	35	43.2	192	2	Q5T035_HUMAN	Q5t035 homo sapien
876	35.5	43.8	929	2	Q5PQW1_RAT	Q5pqw1 rattus norv	949	35	43.2	196	2	Q30587_HELPY	Q30587 helicobacte
877	35.5	43.8	999	2	Q4OHU9_LEIMA	Q4ohu9 leishmania	950	35	43.2	198	2	Q30587_HELPY	Q30587 helicobacte
878	35.5	43.8	1008	2	Q81VQ8_HUMAN	Q81vq8 homo sapien	951	35	43.2	204	2	Q50XN0_ENTHI	Q50xn0 entamoeba h
879	35.5	43.8	1067	2	Q17517_CAEEL	Q17517 caenorhabdi	952	35	43.2	204	2	Q5NT18_ENTHI	Q5nt18 entamoeba h
880	35.5	43.8	1067	2	Q7JMF9_CAEEL	Q7jmf9 caenorhabdi	953	35	43.2	204	2	Q7XK33_ORYSA	Q7xk33 oryza sativ
881	35.5	43.8	1097	1	KPC1_CANAL	P43057 candida alb	954	35	43.2	207	2	Q5BVS9_SCHJA	Q5bvs9 schistosoma
882	35.5	43.8	1097	2	Q5ANK2_CANAL	Q5ank2 candida alb	955	35	43.2	208	2	Q5UC47_HELPY	Q5uc47 helicobacte
883	35.5	43.8	1138	2	Q61ZX3_CAEER	Q61zx3 caenorhabdi	956	35	43.2	209	2	Q8RJF5_HELPY	Q8rjf5 helicobacte
884	35.5	43.8	1149	2	Q6ZPJ4_MOUSE	Q6zpj4 mus musculu	957	35	43.2	210	2	Q30585_HELPY	Q30585 helicobacte
885	35.5	43.8	1151	1	XPO4_HUMAN	Q9cej0 mus musculu	958	35	43.2	210	2	Q5UC40_HELPY	Q5uc40 helicobacte
886	35.5	43.8	1151	1	XPO4_MOUSE	Q9cej0 mus musculu	959	35	43.2	210	2	Q5UC43_HELPY	Q5uc43 helicobacte
887	35.5	43.8	1151	2	Q5VUZ5_HUMAN	Q5vuz5 homo sapien	960	35	43.2	210	2	Q5UC44_HELPY	Q5uc44 helicobacte
888	35.5	43.8	1154	2	Q5ZMR9_CHICK	Q5zmr9 gallus gall	961	35	43.2	210	2	Q5UC48_HELPY	Q5uc48 helicobacte
889	35.5	43.8	1412	2	Q64612_ARATH	Q64612 arabidopsis	962	35	43.2	210	2	Q5UC49_HELPY	Q5uc49 helicobacte
890	35.5	43.8	1506	2	Q6F6B9_HUMAN	Q6f6b9 homo sapien	963	35	43.2	210	2	Q5UC50_HELPY	Q5uc50 helicobacte
891	35.5	43.8	1566	2	Q5VXL9_HUMAN	Q5vxl9 homo sapien	964	35	43.2	210	2	Q5V9R5_HELPY	Q5v9r5 helicobacte
892	35.5	43.8	1566	2	Q4WS60_ASPFU	Q4ws60 aspergillus	965	35	43.2	210	2	Q5V9R6_HELPY	Q5v9r6 helicobacte
893	35.5	43.8	1680	2	Q6MYA4_ASPFU	Q6mya4 aspergillus	966	35	43.2	210	2	Q8RJ56_HELPY	Q8rj56 helicobacte
894	35.5	43.8	1692	2	Q582X4_9TRYP	Q582x4 trypanosoma	967	35	43.2	210	2	Q8RJ83_HELPY	Q8rj83 helicobacte
895	35.5	43.8	2017	2	Q582X4_9TRYP	Q582x4 trypanosoma	968	35	43.2	210	2	Q8RJ83_HELPY	Q8rj83 helicobacte
896	35.5	43.8	4151	2	Q57UX4_9TRYP	Q57ux4 trypanosoma	969	35	43.2	210	2	Q8RJS2_HELPY	Q8rjs2 helicobacte
897	35.5	43.8	9579	2	Q4HWM4_GIBZE	Q4hwm4 gibberella	970	35	43.2	210	2	Q8RJS2_HELPY	Q8rjs2 helicobacte
898	35.5	43.8	10495	2	Q4RE92_TETNG	Q4re92 tetraodon n	971	35	43.2	210	2	Q8VTN5_HELPY	Q8vtn5 helicobacte
899	35	43.2	69	2	Q91TP7_TUHV1	Q91tp7 tupaiid her	972	35	43.2	210	2	Q8VTN7_HELPY	Q8vtn7 helicobacte
900	35	43.2	83	2	Q9PFN0_XYLFA	Q9pfno xylella fas	973	35	43.2	210	2	Q8VTN9_HELPY	Q8vtn9 helicobacte
901	35	43.2	101	2	Q5V0H5_HALMA	Q5v0h5 haloarcula	974	35	43.2	210	2	Q8VTP1_HELPY	Q8vtp1 helicobacte
902	35	43.2	103	2	Q71Z55_LISMF	Q71z55 listeria mo	975	35	43.2	210	2	Q8VTP2_HELPY	Q8vtp2 helicobacte
903	35	43.2	103	2	Q8Y6S1_LISMO	Q8y6s1 listeria mo	976	35	43.2	210	2	Q8VTP3_HELPY	Q8vtp3 helicobacte
904	35	43.2	103	2	Q92B97_LISIN	Q92b97 listeria in	977	35	43.2	210	2	Q8VTP3_HELPY	Q8vtp3 helicobacte
905	35	43.2	103	2	Q8YUR8_ANASP	Q8yur8 anabaena sp	978	35	43.2	210	2	Q8VTP5_HELPY	Q8vtp5 helicobacte
906	35	43.2	106	2	Q4JOF7_AZOVI	Q4jof7 azotobacter	979	35	43.2	210	2	Q8VTP6_HELPY	Q8vtp6 helicobacte
907	35	43.2	107	2	Q4K3V8_PSEF5	Q4k3v8 pseudomonas	980	35	43.2	210	2	Q8VTP7_HELPY	Q8vtp7 helicobacte

981 35 43.2 210 2 Q9FBD1_HELPY Q9fbd1 helicobacte
 982 35 43.2 210 2 Q9FBD2_HELPY Q9fbd2 helicobacte
 983 35 43.2 210 2 Q9FBD3_HELPY Q9fbd3 helicobacte
 984 35 43.2 210 2 Q9FBD4_HELPY Q9fbd4 helicobacte
 985 35 43.2 210 2 Q9FBD5_HELPY Q9fbd5 helicobacte
 986 35 43.2 210 2 Q9FBD6_HELPY Q9fbd6 helicobacte
 987 35 43.2 210 2 Q9FBD7_HELPY Q9fbd7 helicobacte
 988 35 43.2 210 2 Q9FBD8_HELPY Q9fbd8 helicobacte
 989 35 43.2 210 2 Q9FBD9_HELPY Q9fbd9 helicobacte
 990 35 43.2 210 2 Q9FBD10_HELPY Q9fbd10 helicobacte
 991 35 43.2 210 2 Q9FBD11_HELPY Q9fbd11 helicobacte
 992 35 43.2 210 2 Q9FBD12_HELPY Q9fbd12 helicobacte
 993 35 43.2 210 2 Q9FBD13_HELPY Q9fbd13 helicobacte
 994 35 43.2 210 2 Q9FBD14_HELPY Q9fbd14 helicobacte
 995 35 43.2 210 2 Q9FBD15_HELPY Q9fbd15 helicobacte
 996 35 43.2 210 2 Q9FBD16_HELPY Q9fbd16 helicobacte
 997 35 43.2 210 2 Q9FBD17_HELPY Q9fbd17 helicobacte
 998 35 43.2 210 2 Q9FBD18_HELPY Q9fbd18 helicobacte
 999 35 43.2 210 2 Q9FBD19_HELPY Q9fbd19 helicobacte
 1000 35 43.2 210 2 Q9FBD20_HELPY Q9fbd20 helicobacte

ALIGNMENTS

RESULT 1
 Q6CNV5_KLULA PRELIMINARY; PRT; 488 AA.
 ID Q6CNV5_KLULA PRELIMINARY; PRT; 488 AA.
 AC Q6CNV5_KLULA PRELIMINARY; PRT; 488 AA.
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to sp|Q04673 Saccharomyces cerevisiae YLR005W SSL1 TFIIF
 DE subunit.
 GN OrderedLocusNames=KLUA0E09592g;
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15229592; DOI=10.1038/nature02579;
 RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA LaFontaine I., de Montigny J., Marck C., Neugeglie C., Tallia E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin J., Beckerich J.-M., Beyne E., Bleykasten C.,
 RA Boiarane A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.-M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
 RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.-L.;
 RT "Genome evolution in Yeasts."
 RL Nature 430:35-44(2004).
 DR EMBL; CR382125; CAG99471.1; -; Genomic_DNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR GO; GO:0045449; P:regulation of transcription; IEA.
 DR InterPro; IPR004595; Ssl1.
 DR InterPro; IPR007198; Ssl1 like.
 DR InterPro; IPR012170; TFIIF SSL1.
 DR InterPro; IPR002035; VWF A-CH2.
 DR Pfam; PF04056; Ssl1; 1.
 DR PIRSF; PIRSF015919; TFIIF SSL1; 1.
 DR SMART; SM00327; VWA; 1.
 DR TIGRFAMs; TIGR00622; ssl1; 1.
 DR PROSITE; PS00234; VWFA; 1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 488 AA; 54791 MW; E3BD0DC4C55FF08D CRC64;
 Query Match 60.5%; Score 49; DB 2; Length 488;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CSLQEFSLSHGGYVC 14
 Db 360 CACHSTLVHGGYVC 373
 RESULT 2
 Q5S4N1_HUMAN PRELIMINARY; PRT; 362 AA.
 ID Q5S4N1_HUMAN PRELIMINARY; PRT; 362 AA.
 AC Q5S4N1_HUMAN PRELIMINARY; PRT; 362 AA.
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN NUCLEOTIDE SEQUENCE.
 RP Shi Z., Wang H., Feng B., Su G., Huang L.;
 RT "New sequences related to infection with Shigella flexneri 2a.";
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the DEAD box helicase family.
 DR EMBL; AY776161; AAV52794.1; -; mRNA.
 DR Ensembl; ENSG00000124487; Homo sapiens.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000629; DEAD_box.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC_C; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;
 KW Nucleotide-binding.
 FT NON_TER 1 362
 FT NON_TER 362 362
 SQ SEQUENCE 362 AA; 40187 MW; 3EB871DFC708980B CRC64;
 Query Match 59.3%; Score 48; DB 2; Length 362;
 Best Local Similarity 61.5%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 SLQEFSLSHGGYVC 14
 Db 173 SLEDFLYHEGYAC 185
 RESULT 3
 Q9FFN6_ARATH PRELIMINARY; PRT; 365 AA.
 ID Q9FFN6_ARATH PRELIMINARY; PRT; 365 AA.
 AC Q9FFN6_ARATH PRELIMINARY; PRT; 365 AA.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Gb|AA032776.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

```
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RT Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT F1 clones.";
RL DNA Res. 4:215-230(1997).
DR EMBL; AB05234; BAB10471.1; -; Genomic DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 365 AA; 39838 MW; 0E449FDD8753A02 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 365;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLOEFLSHGGVVC 14
|:|:|:|:|
DB 152 CALKFLNHGSTVC 165

RESULT 4
Q8GUG6 ARATH PRELIMINARY; PRT; 367 AA.
AC Q8GUG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein At5g63780.
GN Name=At5g63780;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN NUCLEOTIDE SEQUENCE.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinn P., Tang C.C., Toroumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J., Theologis A.,
RA Davis R.W.;
RA Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RL [2]
RN NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Hsuan V.W., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BT002521; AAC00881.1; -; mRNA.
DR EMBL; BT008456; AAP37815.1; -; mRNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.

KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 367 AA; 40111 MW; 1C846DFEB4B1ACE7 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 367;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSLOEFLSHGGVVC 14
|:|:|:|:|
DB 152 CALKFLNHGSTVC 165

RESULT 5
Q4R9A4 MACFA PRELIMINARY; PRT; 480 AA.
AC Q4R9A4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA clone: QtsA-10416, similar to human DEAD (Asp-Glu-Ala-Asp)
DE box polypeptide 3. X-linked(DDX3X), transcript variant 1.
DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheidae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RA "DNA sequences of macaque genes expressed in brain or testis and its
RT evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AB168192; BAE00317.1; -; mRNA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR SMART; SM00487; Helicase_C; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
SQ SEQUENCE 480 AA; 52923 MW; CDSF3F1F7557A809 CRC64;

Query Match 59.3%; Score 48; DB 2; Length 480;
Best Local Similarity 61.5%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGVVC 14
|:|:|:|:|
DB 289 SLEDFLYHEGYAC 301

RESULT 6
DDX3Y PONPY STANDARD; PRT; 658 AA.
ID DDX3Y PONPY STANDARD; PRT; 658 AA.
AC Q5RF43;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN Name=DDX3Y;
OS Pongo pygmaeus (Orangutan).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RG The German cDNA consortium;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Probable ATP-dependent RNA helicase. May play a role in
CC spermatogenesis (By similarity).
CC -1- SUBCELLULAR LOCATION: Shuttles between the nucleus and the
CC cytoplasm in an XPO1-dependent manner (By similarity).
CC -1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CR857318; CAH89614.1; -; mRNA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR00629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; DNA-binding; Helicase; Hydrolase; Nuclear protein;
FT Nucleotide-binding; RNA-binding.
FT NP BIND 222 229 ATP (Potential).
FT MOTIF 345 348 DEAD box.
FT SEQUENCE 658 AA; 73043 MW; 21237B600A5FDD6D CRC64;
SQ
Query Match 59.3%; Score 48; DB 1; Length 658;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 2 SLQEFSLSHGGYVC 14
Db 454 SLEDFLYHEGYAC 466
RESULT 7
DDX3Y_HUMAN STANDARD; PRT; 660 AA.
AC O15523; Q8IYV7;
DT 15-JUL-1999 (Rel. 38, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN Name=DDX3Y; Synonyms=DBY;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1] NUCLEOTIDE SEQUENCE [MRNA].
RP MEDLINE=98022381; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Lahn B.T., Page D.C.;
RT "Functional coherence of the human Y chromosome.";
RL Science 278:675-680 (1997).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RN SUBCELLULAR LOCATION.
RP PubMed=15383328; DOI=10.1016/j.yexor.2004.07.005;
RX Sekiguchi T., Iida H., Fukumura J., Nishimoto T.;
RT "Human DDX3Y, the Y-encoded isoform of RNA helicase DDX3, rescues a
RT hamster temperature-sensitive Etr24 mutant cell line with a DDX3X
RT mutation.";
RL Exp. Cell Res. 300:213-222 (2004).
RN [4]
RN TISSUE SPECIFICITY.
RP PubMed=15294876; DOI=10.1093/hmg/ddh240;
RX Dittion H.J., Zimmer J., Kamp C., Rajpert-De Meyts E., Vogt P.H.;
RT "The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is
RT limited to the male germ cells by translation control.";
RL Hum. Mol. Genet. 13:2333-2341 (2004).
CC -1- FUNCTION: Probable ATP-dependent RNA helicase. May play a role in
CC spermatogenesis.
CC -1- SUBCELLULAR LOCATION: Shuttles between the nucleus and the
CC cytoplasm in an XPO1-dependent manner.
CC -1- TISSUE SPECIFICITY: Testis-specific. Expressed predominantly in
CC spermatogonia.
CC -1- DISEASE: DDX3Y is deleted in severe nonobstructive
CC hypogonadism [MIM:400042].
CC -1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AF000985; AAC51832.1; -; mRNA.
DR EMBL; AF000984; AAC51831.1; -; mRNA.
DR EMBL; BC034942; AAH34942.1; -; mRNA.
DR HSSP; P10081; 1FUK.
DR Ensembl; ENSG00000067048; Homo sapiens.
DR HGNC; HGNC:2699; DDX3Y.
DR MIM; 400010; --.
DR MIM; 400042; --.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW ATP-binding; DNA-binding; Helicase; Hydrolase; Nuclear protein;
KW Nucleotide-binding; RNA-binding.
FT NP BIND 222 229 ATP (Potential).
FT MOTIF 345 348 DEAD box.
FT CONFLICT 46 46 R -> K (in Ref. 1).
FT CONFLICT 219 219 M -> V (in Ref. 1).
FT CONFLICT 628 628 N -> D (in Ref. 1).
SQ SEQUENCE 660 AA; 73154 MW; 0C370E9367952AB6 CRC64;
Query Match 59.3%; Score 48; DB 1; Length 660;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 2 SLQEFSLSHGGYVC 14
```

```

Db      454 SLEDFLYHEGYAC 466
||||:|||||
RESULT 8
DDX3V PANTR STANDARD; PRT; 660 AA.
AC O6GVME;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN Name=DDX3Y;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pan.
OC NCBI_TaxID=9598;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RC MEDLINE=89249320; PubMed=2720782; DOI=10.1016/0092-8674(89)90125-6;
RX Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Putative ATP-dependent RNA helicase. Possible role in a
CC key step of the spermatogenic process.
CC -!- TISSUE SPECIFICITY: Testis.
CC -!- DEVELOPMENTAL STAGE: High levels of PL10 during the meiotic and
CC haploid stages of spermatogenesis.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL; J04847; AAA39942.1; -; mRNA.
CC EMBL; AK029542; BAC26505.1; -; mRNA.
CC PIR; A32378; A32378.
CC HSSP; Q58083; 1HV8.
CC Ensembl; ENSMUSG00000039224; Mus musculus.
CC MGI; MGI:91842; D1Pae1.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR011545; DEAD/DEAH_N.
CC InterPro; IPR000629; DEAD_box.
CC InterPro; IPR001650; Helicase_C.
CC Pfam; PF00270; DEAD; 1.
CC Pfam; PF00271; Helicase_C; 1.
CC PROSITE; PS00039; DEAD ATP HELICASE; 1.
CC Name=Blfasi; Synonyms=PL10;
OS Mus musculus (Mouse).

KW
Query Match 59.3%; Score 48; DB 1; Length 660;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQERLSHGYYC 14
||||:|||||
Db      454 SLEDFLYHEGYAC 466
||||:|||||
RESULT 9
PL10_MOUSE STANDARD; PRT; 660 AA.
AC P16381;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Putative ATP-dependent RNA helicase PL10.
GN Name=Blfasi; Synonyms=PL10;
OS Mus musculus (Mouse).

```

KW Helicase; Hydrolase; Nucleotide-binding; RNA-binding; Spermatogenesis.
 FT NP_BIND 223 230 ATP (Potential).
 FT DNA_BIND 495 514 Potential.
 FT MOTIF 346 349 DEAD box.
 FT COMPBIAS 581 660 Gly/Ser-rich.
 SQ SEQUENCE 660 AA; 73141 MW; 50AD4B6A131AE603 CRC64;
 Query Match 59.3%; Score 48; DB 1; Length 660;
 Best Local Similarity 61.5%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 SLOEFLSHGGYVC 14
 Db 455 SLEDFLYHEGYAC 467
 RESULT 10
 DDX3X HUMAN
 ID DDX3X_HUMAN STANDARD; PRT; 661 AA.
 AC O00571; O15536;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE DEAD-box protein 3, X-chromosomal (Helicase-like protein 2) (HLP2)
 DE (DEAD-box, X isoform).
 GN Name=DDX3X; Synonyms=DBX, DDX3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RP [1]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Hippocampus, and Liver;
 RA Chung J., Lee S.-G., Song K.;
 RT "Identification of a human homolog of a putative RNA helicase gene
 (mDEAD3) expressed in mouse erythroid cells.";
 RL Korean J. Biochem. 27:193-197(1995).
 [2]
 RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, SUBCELLULAR LOCATION, AND
 RP INTERACTION WITH HCV CORE PROTEIN.
 RC TISSUE=Liver;
 RX MEDLINE=99263161; PubMed=10329544; DOI=10.1006/viro.1999.9659;
 RA Owsianka A.M., Patel A.H.;
 RT "Hepatitis C virus core protein interacts with a human DEAD box
 protein DDX3.";
 RL Virology 257:330-340(1999).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98022381; PubMed=9381176; DOI=10.1126/science.278.5338.675;
 RA Lahn B.T., Page D.C.;
 RT "Functional coherence of the human Y chromosome.";
 RL Science 278:675-680(1997).
 [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J.J., Helton E., Kettelman M., Maman A., Rodrigues S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [5]
 RP PROTEIN SEQUENCE OF 1-9, AND ACETYLATION.
 RX MEDLINE=20318637; PubMed=10859333; DOI=10.1084/jem.191.12.2083;
 RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,
 RA Lopez de Castro J.A.;
 RT "An N-acetylated natural ligand of human histocompatibility leukocyte
 antigen (HLA)-B39. Classical major histocompatibility complex class I
 proteins bind peptides with a blocked NH(2) terminus in vivo.";
 RL J. Exp. Med. 191:2083-2092(2000).
 [6]
 RP FUNCTION IN HIV-1 RNA EXPORT AND REPLICATION, IDENTIFICATION IN A
 RP COMPLEX WITH XPO1 AND REV, INTERACTION WITH XPO1, MUTAGENESIS OF
 RP LYS-229 AND SER-381, AND SUBCELLULAR LOCATION.
 RX PubMed=15507209; DOI=10.1016/j.cell.2004.09.029;
 RA Yedavalli V.S., Neuveut C., Chi Y.-H., Kleiman L., Jeang K.-T.;
 RT "Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export
 function.";
 RL Cell 119:381-392(2004).
 CC -1- FUNCTION: ATP-dependent RNA helicase. Acts as a cofactor for XPO1-
 CC mediated nuclear export of incompletely spliced HIV-1 Rev RNAs.
 CC Also involved in HIV-1 replication. Interacts specifically with
 CC hepatitis C virus core protein resulting in a change in
 CC intracellular location.
 CC -1- SUBUNIT: Found in a complex with Rev and XPO1. Interacts with
 CC XPO1. Interacts with HCV core protein.
 CC -1- INTERACTION:
 CC Q00653:NFKB2; NbExp=1; IntAct=EBI-353779; EBI-307326;
 CC -1- SUBCELLULAR LOCATION: Located predominantly in nuclear speckles
 CC and, at low levels, throughout the cytoplasm. Located to the outer
 CC side of nuclear pore complexes (NPC). Shuttles between the nucleus
 CC and the cytoplasm in a XPO1-dependent manner.
 CC -1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
 CC subfamily.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; U50553; AAB95637.1; -; mRNA.
 CC EMBL; AF061337; AAC34298.1; -; mRNA.
 CC EMBL; AF000983; AAC51830.1; -; mRNA.
 CC EMBL; AF000982; AAC51829.1; -; mRNA.
 CC EMBL; BC011819; AAH11819.1; -; mRNA.
 CC HSP; Q58083; IHV8.
 CC InAct; Q00571; -;
 CC SWISS-2DPAGE; O00571; HUMAN.
 CC Ensembl; ENSG00000124487; Homo sapiens.
 CC HGNC; HGNC:2745; DDX3X.
 CC H-InvDB; HIX0016737; -;
 CC MIM; 300160; -;
 CC GO; GO:0005737; C:cytoplasm; IDA.
 CC GO; GO:0005634; C:nucleus; IDA.
 CC GO; GO:004004; F:ATP-dependent RNA helicase activity; TAS.
 CC InterPro; IPR001410; DEAD.
 CC InterPro; IPR011545; DEAD/DEAH_N.
 CC InterPro; IPR000629; DEAD_box.
 CC InterPro; IPR001650; Helicase_C.
 CC Pfam; PF00270; DEAD; 1.
 CC Pfam; PF00271; Helicase_C; 1.
 CC SMART; SM00487; DEXDc; 1.
 CC SMART; SM00490; HELICc; 1.
 CC PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 KW Acetylation; ATP-binding; Direct protein sequencing; DNA-binding;
 KW Helicase; Hydrolase; Nuclear protein; Nucleotide-binding;
 KW Phosphorylation; RNA-binding.
 FT INIT MET 0
 FT NP_BIND 223 230 ATP (Potential).

and, at low levels, throughout the cytoplasm. Located to the outer side of nuclear pore complexes (NPC). Shuttles between the nucleus and the cytoplasm in a XPO1-dependent manner (By similarity).

-1- TISSUE SPECIFICITY: Developmentally regulated.

-1- DEVELOPMENTAL STAGE: Expressed in oocytes. Ubiquitously found in 9 days post-conception embryo, at later stages it is restricted to brain and kidney.

-1- SIMILARITY: Belongs to the DEAD box helicase family. DDX3 subfamily.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; Z38117; CAA86261.1; -; mRNA.
EMBL; L25126; AAS3630.1; -; mRNA.
PIR; I84741; I84741.
HSSP; Q58083; LHV8.
Ensembl; ENSMUSG0000000787; Mus musculus.
MGI; MGI:103064; Ddx3x.
GO; GO:0005737; C:cytoplasm; ISS.
GO; GO:0005634; C:nucleus; ISS.
GO; GO:0004004; F:ATP-dependent RNA helicase activity; ISS.
InterPro; IPR001410; DEAD.
InterPro; IPR011545; DEAD/DEAH_N.
InterPro; IPR000629; DEAD_box.
InterPro; IPR001650; Helicase_C.
Pfam; PF00270; DEAD; 1.
Pfam; PF00271; Helicase_C; 1.
PROSITE; PS00039; DEAD ATP HELICASE; 1.
Acetylation; ATP-binding; Direct protein sequencing; DNA-binding;
Helicase; Hydrolase; Nuclear protein; Nucleotide-binding;
Phosphorylation; RNA-binding.
INIT_MET 0
NP_BIND 223 230 ATP (Potential).
REGION 259 516 Necessary for interaction with XPO1 (By similarity).
MOTIF 346 349 DEAD box.
COMPBIAS 581 661 Gly/Ser-rich.
COMPBIAS 608 615 Poly-Ser.
COMPBIAS 623 629 Poly-Gly.
COMPBIAS 632 640 Poly-Gly.
MOD_RES 1 1 N-acetyls erine.
MOD_RES 73 73 Phosphoserine.
SEQUENCE 661 AA; 72970 MW; A1E1FA85D19F57B CRC64;
Query Match 59.3%; Score 48; DB 1; Length 661;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLOEFLSHGGYVC 14
||:|||||
Db 455 SLEDFLYHEGYAC 467

RESULT 12
Q5J5I3 HUMAN PRELIMINARY; PRT; 662 AA.
AC Q5J5I3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked.
GN Name=DDX3X; ORFNames=RP1-16915.3-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.

FT REGION 259 516 Necessary for interaction with XPO1.
FT MOTIF 346 349 DEAD box.
FT COMPBIAS 581 661 Gly/Ser-rich.
FT MOD_RES 1 1 N-acetyls erine.
FT MOD_RES 73 73 Poly-Ser.
FT MUTAGEN 229 229 K->E: Abolishes ATPase activity and RNA-unwinding activity.
FT MUTAGEN 381 381 S->L: Abolishes ATPase activity and RNA-unwinding activity.
FT CONFLICT 49 49 K -> R (in Ref. 3).
SQ SEQUENCE 661 AA; 73112 MW; F0F03DD8FBC00A65 CRC64;
Query Match 59.3%; Score 48; DB 1; Length 661;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLOEFLSHGGYVC 14
||:|||||
Db 455 SLEDFLYHEGYAC 467

RESULT 11
DDX3X MOUSE STANDARD; PRT; 661 AA.
AC Q62167; O09060; O09143;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DEAD-box protein 3, X-chromosomal (DEAD-box RNA helicase DEAD3)
DE (mDEAD3) (Embryonic RNA helicase) (DIPas1-related sequence 2).
GN Name=Ddx3x; Synonyms=DIPas1-rs2, Ddx3, Dead3, Erh;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6, and DBA;
RX MEDLINE=94192995; PubMed=8948440;
RA Sowden J.C., Putt W., Morrison K., Beddington R., Edwards Y.;
RA "The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of RNA helicases";
RL Biochem. J. 308:839-846 (1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Erythrocyte;
RX MEDLINE=94192995; PubMed=8144024; DOI=10.1016/0378-1119(94)90541-X;
RA Gee S.L., Conboy J.G.;
RA "Mouse erythroid cells express multiple putative RNA helicase genes exhibiting high sequence conservation from yeast to mammals."; Gene 140:171-177 (1994).
RN [3]
RP PROTEIN SEQUENCE OF 1-9, AND ACETYLATION.
RX MEDLINE=20318637; PubMed=10859333; DOI=10.1084/jem.191.12.2083;
RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J., Lopez de Castro J.A.;
RA "An N-acetylated natural ligand of human histocompatibility leukocyte antigen (HLA)-B39. Classical major histocompatibility complex class I proteins bind peptides with a blocked NH(2) terminus in vivo."; J. Exp. Med. 191:2083-2092 (2000).
RN [4]
RP PHOSPHORYLATION SITE SER-73.
RX PubMed=14729942; DOI=10.1074/mcp.D300003-MCP200;
RA Shu H., Chen S., Bi Q., Mumby M., Brecken D.L.;
RA "Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line."; Mol. Cell. Proteomics 3:279-286 (2004).
CC -1- FUNCTION: Putative ATP-dependent RNA helicase. It may play a role in translational activation of mRNA in the oocyte and early embryo.
CC -1- SUBUNIT: Interacts with XPO1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Located predominantly in nuclear speckles

Brown A.;
Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: Belongs to the DEAD box helicase family.
EMBL; AL391647; CAI41416.1; -; Genomic DNA.
Ensembl; ENSG00000124487; Homo sapiens.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003676; F:nucleic acid binding; IEA.
InterPro; IPR001410; DEAD.
InterPro; IPR011545; DEAD/DEAH_N.
InterPro; IPR000629; DEAD box_
InterPro; IPR001650; Helicase_C.
Pfam; PF00270; DEAD; 1.
Pfam; PF00271; Helicase_C; 1.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELIC; 1.
PROSITE; PS00039; DEAD ATP HELICASE; 1.
ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
SEQUENCE 662 AA; 73243 MW; 7074D2B8A6EBBF09 CRC64;
Query Match 59.3%; Score 48; DB 2; Length 662;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
Db 456 SLEDFLYHEGYAC 468
||:|||||

RESULT 13
Q8K5DS MESAU PRELIMINARY; PRT; 662 AA.
AC Q8K5DS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA helicase.
GN Name=BX;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Cricetidae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825207; PubMed=12944373; DOI=10.1093/jb/mvg126;
RA Fukumura J., Noguchi E., Sekiguchi T., Nishimoto T.;
RT "A temperature-sensitive mutant of the mammalian RNA helicase, DEAD-
BOX X isoform, DBX, defective in the transition from G1 to S phase.";
RL J. Biochem. 134:71-82(2003).
CC -1- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AB080116; BAB91216.1; -; mRNA.
DR HSSP; Q58083; 1HV8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD box_
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
SEQUENCE 662 AA; 73210 MW; 688A243DA6D0CEE2 CRC64;
Query Match 59.3%; Score 48; DB 2; Length 662;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
Db 456 SLEDFLYHEGYAC 468
||:|||||

RESULT 14
Q59GX6 HUMAN PRELIMINARY; PRT; 674 AA.
AC Q59GX6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AB208983; BAD92220.1; -; mRNA.
DR Ensembl; ENSG00000124487; Homo sapiens.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD box_
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
ATP-binding; Helicase; Hydrolase; Nucleotide-binding.
FT NON_TER 1
SQ SEQUENCE 674 AA; 74579 MW; F1E757F913FC64DF CRC64;
Query Match 59.3%; Score 48; DB 2; Length 674;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
Db 468 SLEDFLYHEGYAC 480
||:|||||

RESULT 15
Q8QH15 TYRXY PRELIMINARY; PRT; 202 AA.
AC Q8QH15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oocyte maturation factor Mos (Fragment).
GN Name=c-mos;
OS Tyrannus tyrannus (Eastern kingbird).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Tyrannidae; Tyrannus.
OX NCBI_TaxID=43165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21829499; PubMed=11839199; DOI=10.1098/rspb.2001.1883;
RA Barker F.K., Barrowclough G.F., Groth J.G.;
RT "A phylogenetic hypothesis for passerine birds: taxonomic and

RT biogeographic implications of an analysis of nuclear DNA sequence
 RT data.", B, Biol. Sci. 269:295-308(2002).
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.

DR EMBL; AY056971; AAL18088.1; -; Genomic_DNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.

DR Pfam; PF00069; PKinase; 1.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Nucleotide-Binding;
 KW Serine/threonine-protein kinase; Transferase.

FT NON_TER 1 202
 FT NON_TER 202 202

SQ SEQUENCE 202 AA; 21898 MW; BFC6D57F86680A9D CRC64;

Query Match 56.8%; Score 46; DB 2; Length 202;
 Best Local Similarity 62.5%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 1 CS--LOEFLSHGGYVC 14

DB 164 CSQKLEGLSHSSYVC 179

RESULT 16

ID Q7XK97 ORYSA

AC Q7XK97 ORYSA PRELIMINARY; PRT; 358 AA.

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE OSUNB0020J19.13 protein.

GN Name=OSUNB0020J19.13;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;

RP NUCLEOTIDE SEQUENCE.

RA MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;

RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,

RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,

RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.-S., Yu Z., Fan D.,

RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,

RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,

RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,

RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,

RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,

RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,

RA Han B.;

RT "Sequence and analysis of rice chromosome 4.";

RL Nature 420:316-320(2002).

DR EMBL; AL606556; CA05784.2; -; Genomic_DNA.

DR Gramine; Q7XK97; -.

DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.

DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0016567; P:protein ubiquitination; IEA.

DR InterPro; IPR001309; ICE_P20.

DR InterPro; IPR001841; Znf_RING.

DR Pfam; PF00097; zf-C3HC4; 1.

DR PROSITE; PS01121; CASPASE_HIS; UNKNOWN 1.

KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.

SQ SEQUENCE 358 AA; 38844 MW; F89A0D01AC61771D CRC64;

Query Match 56.8%; Score 46; DB 2; Length 358;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 CSLOEFLSHGGYVC 14

DB 120 CALKWFITHGSTVC 133

RESULT 17

ID Q69BL0 MANSE

AC Q69BL0 MANSE PRELIMINARY; PRT; 666 AA.

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Pattern recognition serine proteinase precursor.

GN Name=PRSP;

OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;

OC Spingidae; Sphinginae; Manduca.

OX NCBI_TaxID=7130;

RP NUCLEOTIDE SEQUENCE.

RA PubMed=15190055; DOI=10.1074/jbc.M404584200;

RA Ji C., Wang Y., Guo X., Hartson S., Jiang H.;

RT "A pattern recognition serine proteinase triggers the prophenoloxidase

activation cascade in the tobacco hornworm, *Manduca sexta*.";

RL J. Biol. Chem. 279:34101-34106(2004).

DR EMBL; AY380790; AAR29602.1; -; mRNA.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR002172; LDL_receptor_A.

DR InterPro; IPR001314; Peptidase_S1A.

DR InterPro; IPR001254; Peptidase_S1_S6.

DR InterPro; IPR000436; Sushi_SCR_CCP.

DR Pfam; PF00057; Ldl_recept_a; 5.

DR Pfam; PF00084; Sushi; 1.

DR Pfam; PF00089; Trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00261; LDLRECEPTOR.

DR SMART; SM00032; CCP; 1.

DR SMART; SM00192; LDLa; 5.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS01209; LDLRA_1; 4.

DR PROSITE; PS00068; LDLRA_2; 3.

DR PROSITE; PS50923; SUSHI; 1.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

FT SIGNAL.

FT CHAIN 1 17 Potential.

FT SIGNAL 18 666 pattern recognition serine proteinase.

SQ SEQUENCE 666 AA; 73670 MW; 880439BC41B9238F CRC64;

Query Match 56.8%; Score 46; DB 2; Length 666;
 Best Local Similarity 61.5%; Pred. No. 49;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CSLOEFLSHGGYVC 13

DB 258 CVLPYFPHGGYV 270

RESULT 18

ID Q93ZD2 ARATH

AC Q93ZD2 ARATH PRELIMINARY; PRT; 363 AA.

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE AT5G63780/MBK5 26.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowler L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlins-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Mirananda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057621; AAL14116.1; -; mRNA.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00744; RINGv; 1.
DR KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
KW SEQUENCE 363 AA; 39458 MW; 8BA2C124FB76BE8E CRC64;
Query Match 55.6%; Score 45; DB 2; Length 363;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Oy 1 CSIQEFLSHGGYVC 14
Db 140 CALKWFVNHGTVTC 153
RESULT 19
O9CSA2 ARATH PRELIMINARY; PRT; 363 AA.
AC O9CSA2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein At5g08750 (Hypothetical protein
DE At5g08750/T2K12.100).
GN Name=At5g08750; Synonyms=At5g08750/T2K12.100;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL590346; BAC35880.1; -; mRNA.
DR EMBL; AK117992; BAC42627.1; -; mRNA.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR011015; RINGv.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00744; RINGv; 1.
DR KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
KW SEQUENCE 363 AA; 39430 MW; 6B1A8E05A45F7281 CRC64;
Query Match 55.6%; Score 45; DB 2; Length 363;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Oy 1 CSIQEFLSHGGYVC 14
Db 140 CALKWFVNHGTVTC 153
RESULT 20
QSER66 MONAL PRELIMINARY; PRT; 376 AA.
AC QSER66;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DEAD box RNA helicase-PL10A (Fragment).
OS Monopterus albus (Swamp eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Synbranchiiformes;
OC Synbranchidae; Monopterus.
OX NCBI_TaxID=43700;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Maoyu P., Ping S.;
RT "Cloning and Sequence Analysis of Monopterus albus DEAD-box Family
RT Gene Pl10.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; AY786203; AAW78518.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding.
FT NON_TER 1
FT NON_TER 376
SQ SEQUENCE 376 AA; 42248 MW; B42BA83E0B4BF859 CRC64;
Query Match 55.6%; Score 45; DB 2; Length 376;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Oy 2 SLOEFLSHGGYVC 14
Db 304 ALDFLYHEGVAC 316
RESULT 21
Q518T7 ENTHI PRELIMINARY; PRT; 448 AA.
ID Q518T7 ENTHI
AC Q518T7;

```
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 5'-3' exonuclease, putative.
GN ORFNames=45.t00020;
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN NCU
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinovitch E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000190; EAL49366.1; -; Genomic DNA.
DR GO; GO:0004527; F:exonuclease activity; IEA.
DR InterPro; IPR012282; Cycochrome_c.R.
DR InterPro; IPR012351; Cycokine_4_hlx.
DR InterPro; IPR004859; Puc_53exo.
DR InterPro; IPR012335; Thiorodoxin_fold.
DR Pfam; PF03159; XRN_N; 1.
KW Exonuclease.
SQ SEQUENCE 448 AA; 52207 MW; C860798270121D04 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 448;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QEFSLSHGGV 13
:|||||:
DB 345 REFLPHGGYI 354

RESULT 22
RS1_CHLMU STANDARD; PRT; 570 AA.
AC P38016;
DT 01-OCT-1994 (Rel. 30, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 30S ribosomal protein S1 (70 kDa antigen).
GN Name=rpS4; OrderedLocustNames=TC0373;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN NCU
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MoPn / N199;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN NCU
RP NUCLEOTIDE SEQUENCE OF 1-180.
```

```
RC STRAIN=MoPn;
RX MEDLINE=89123039; PubMed=2644193;
RA Sardinia L.M., Engel J.N., Ganem D.;
RT "Chlamydial gene encoding a 70-kilodalton antigen in Escherichia coli:
RT analysis of expression signals and identification of the gene
RT product.";
RL J. Bacteriol. 171:335-341(1989).
CC -!- FUNCTION: Binds mRNA; thus facilitating recognition of the
CC initiation point. It is needed to translate mRNA with a short
CC Shine-Dalgarno (SD) purine-rich sequence (By similarity).
CC -!- SIMILARITY: Belongs to the ribosomal protein S1P family.
CC -!- SIMILARITY: Contains 6 S1 motif domains.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE002304; AAF39231.1; ALT_INIT; Genomic_DNA.
DR EMBL; M23000; AAA23167.1; -; Genomic_DNA.
DR PIR; A32246; A32246.
DR PIR; A81710; A81710.
DR HSPP; P05055; ISRO.
DR Siena-2DPAGE; P38016; -.
DR TIGR; TC0373; -.
DR InterPro; IPR012340; OB_NA_bd_sub.
DR InterPro; IPR000110; Ribosomal_S1.
DR InterPro; IPR003029; S1_RNA_bd.
DR Pfam; PF00575; S1; 6.
DR PRINTS; PR00681; RIBOSOMALS1.
DR SMART; SM00316; S1; 6.
DR TIGRFAMs; TIGR00717; rpsA; 1.
DR PROSITE; PS50126; S1; 6.
KW Complete proteome; Repeat; Ribonucleoprotein; Ribosomal protein;
KW RNA-binding.
FT DOMAIN 52 116 S1 motif 1.
FT DOMAIN 134 199 S1 motif 2.
FT DOMAIN 220 288 S1 motif 3.
FT DOMAIN 305 375 S1 motif 4.
FT DOMAIN 392 462 S1 motif 5.
FT DOMAIN 479 548 S1 motif 6.
FT CONFLICT 139 139 G -> A (in Ref. 2).
FT CONFLICT 143 143 R -> P (in Ref. 2).
FT CONFLICT 177 180 GKVC -> RESL (in Ref. 2).
SQ SEQUENCE 570 AA; 63610 MW; 5715B2C711518A24 CRC64;

Query Match 55.6%; Score 45; DB 1; Length 570;
Best Local Similarity 63.6%; Pred. No. 62;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLOEFLSHGGY 12
|:|||||:
DB 546 SIXEFLAHGGH 556

RESULT 23
Q4KXN1_MOUSE PRELIMINARY; PRT; 611 AA.
ID Q4KXN1_MOUSE PRELIMINARY;
AC Q4KXN1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE 1100001A2Rik protein.
GN Name=1100001A2Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCU
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6Ncr; TISSUE=Hematopoietic Stem Cell;
```


[4]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Head;
 RC MEDLINE=20499374; PubMed=1042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RL prepare full-length cDNA libraries for rapid discovery of new genes.";
 RN Genome Res. 10:1617-1630(2000).
 [5]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Head;
 RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Azawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuura T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Head;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu M., Hiramoto K., Hiraoka T., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi P., Takaku-Akai H., Tanaka T., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK030661; BAC27067.1; -; mRNA.
 DR Ensembl; ENSMUSG0000032567; Mus musculus.
 DR MGI; MGI:1913845; 110001A21Rik.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0004518; F:nucleic acid binding; IEA.
 DR InterPro; IPR000513; Exo_N1.
 SQ SEQUENCE 672 AA; 76370 MW; 51F219D759443543 CRC64;
 Query Match 55.6%; Score 45; DB 2; Length 672;
 Best Local Similarity 58.3%; Pred. No. 74;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LQEFSLSHGGYVC 14
 Db 304 LQDFQVGSVC 315
 RESULT 26
 Q52L23_XENLA PRELIMINARY; PRT; 695 AA.
 AC Q52L23;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Oocytes;
 RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Oocytes;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Beasak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Oocytes;
 RC Klein S., Gerhardt D.S.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 CC -1; SIMILARITY: Belongs to the DEAD box helicase family.
 DR EMBL; BC094037; AAH94097.1; -; mRNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000629; DEAD_box_
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
 KW ATP-binding; Helicase; Hydrolase; Hypothetical protein;
 KW Nucleotide-binding.
 SQ SEQUENCE 695 AA; 77045 MW; DC5C235C3C06689A CRC64;
 Query Match 55.6%; Score 45; DB 2; Length 695;
 Best Local Similarity 53.8%; Pred. No. 77;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 SLQEFSLSHGGYVC 14
 Db 495 ALQDFLYHEGYAC 507
 RESULT 27
 AN3_XENLA
 ID AN3_XENLA STANDARD; PRT; 697 AA.
 AC F24346;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Putative ATP-dependent RNA helicase An3.
 DE Name=AN3;
 GN Xenopus laevis (African clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.

```

OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91141586; PubMed=1996140; DOI=10.1038/349717a0;
RA Gururajan R., Perry-O'Keefe H., Melton D.A., Weeks D.L.;
RT "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
RL RNA helicase.";
RL Nature 349:717-719 (1991).
CC -!- FUNCTION: Putative ATP-dependent RNA helicase.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DDX3
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X57328; CAA40605.1; -; mRNA.
DR PIR; S13654; S13654.
DR HSP; P10081.1; FUK.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding; RNA-binding.
FT NP_BIND 265 272 ATP (By similarity).
FT MOTIF 388 391 DEAD box.
FT COMPIAS 623 697 Gly/Ser-rich.
SQ SEQUENCE 697 AA; 77303 MW; F3DD23EB60B2E2BF CRC64;

Query Match 55.6%; Score 45; DB 1; Length 697;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
:|:|:|:|:|
Db 497 ALEDFLYHEGYAC 509

RESULT 28
Q7ZXJ0_XENLA PRELIMINARY; PRT; 697 AA.
ID Q7ZXJ0_XENLA PRELIMINARY; PRT; 697 AA.
AC Q7ZXJ0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P110-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Statchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Bosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.;
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.
DR EMBL; BC044972; AAH44972.1; -; mRNA.
DR HSP; OS8083; IHV8.
DR GO; GO:0005524; F-ATP binding; IEA.
DR GO; GO:0008026; F-ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F-hydrolase activity; IEA.
DR GO; GO:0003676; F-nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW ATP-binding; Helicase; Hydrolase; Nucleotide-binding.
SQ SEQUENCE 697 AA; 77133 MW; C6ED241A0024A384 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 697;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SLOEFLSHGGYVC 14
:|:|:|:|:|
Db 497 ALEDFLYHEGYAC 509

RESULT 29
Q6P4J3_XENTR PRELIMINARY; PRT; 699 AA.
ID Q6P4J3_XENTR PRELIMINARY; PRT; 699 AA.
AC Q6P4J3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC76021.
GN Name=MGC76021;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

```


RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichert-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Peri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -I- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAFB01000025; EAL51687.1; -; Genomic_DNA.
KW Exonuclease.
SQ SEQUENCE 948 AA; 111242 MW; D1BB9F8186A8034C CRC64;

Query Match 55.6%; Score 45; DB 2; Length 948;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QEFLSHGGYV 13
DB 386 REFLPGGYI 395
:|||||

RESULT 31
Q9FJE2 ARATH PRELIMINARY; PRT; 651 AA.
AC Q9FJE2
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tall-like non-LTR retroelement protein-like.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:297-308(1998).
DR EMBL; AB015475; BAB08362.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011424; Cl 3.
DR InterPro; IPR002219; DAG PE-bind.
DR InterPro; IPR004146; DCL.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF03107; Cl 2; 2.
DR Pfam; PF07649; Cl 3; 3.
DR SMART; SM00109; Cl 1; 2.
DR SMART; SM00249; PHD; 4.
SQ SEQUENCE 651 AA; 72730 MW; AAFDD67F0D8B869A CRC64;

Query Match 54.9%; Score 44.5; DB 2; Length 651;
Best Local Similarity 50.0%; Pred. No. 87;
Matches 7; Conservative 5; Mismatches 1; Indels 1; Gaps 1

QY 1 CSLQEFSLSHGGYVC 14
DB 29 CGVDFV-YGGYVC 41
|:::|::|

RESULT 32
Q5DELO SCHJA
ID Q5DELO SCHJA PRELIMINARY; PRT; 266 AA.

AC	Q5DEL0;
DT	10-MAY-2005 (TrEMBLrel. 30, Created)
DT	10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT	10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE	Hypothetical protein.
OS	Schistosoma japonicum (Blood fluke).
OC	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC	Schistosomatidae; Schistosomatidae; Schistosomata; Schistosoma.
NCBI_TaxID=6182;	[1]
RN	NCBI TaxID=6182;
RP	NUCLEOTIDE SEQUENCE.
RA	Han Z.,
RT	The full-length cDNA sequences of Schistosoma japonicum genes.";
RL	Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY814014; AAW25746.1; -, mRNA.
KW	Hypothetical protein.
SQ	SEQUENCE 266 AA; 29953 MW; E81E6C772947038E CRC64;
Query Match	54.3%; Score 44; DB 2; Length 266;
Best Local Similarity	58.3%; Pred. No. 40;
Matches	7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy	3 LQEFLSHGGYVC 14 : :
Dd	96 LLEFMHGGVLC 107 : :
RESULT 33	
ID	081633 ARATH PRELIMINARY; PRT; 319 AA.
AC	081633;
DT	01-NOV-1998 (TrEMBLrel. 08, Created)
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Inositol 1,3,4-trisphosphate 5/6-kinase.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;	[1]
RN	NCBI TaxID=3702;
RP	NUCLEOTIDE SEQUENCE.
RA	MEDLINE=97271384; PubMed=9126335; DOI=10.1006/bbrc.1997.6355;
RL	Wilson M.P.; Majerus P.W.J
RT	"Characterization of a cDNA encoding Arabidopsis thaliana inositol
RT	1,3,4-trisphosphate 5/6-kinase.";
RL	Biochem. Biophys. Res. Commun. 232:678-681(1997).
[2]	
RN	NUCLEOTIDE SEQUENCE.
RA	Wilson M.P.; Majerus P.W.J
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF080173; AAC28859.1; -, mRNA.
PIR	JC5401; JCS401.
GO	GO:0016301; F.kinase activity; IEA.
GO	GO:0016874; F.ligase activity; IEA.
InterPro	IPR011761; ATP GRASP.
InterPro	IPR008656; Insi34_P3_kin.
Pfam	PF05770; Insi34_P3_kin; 1.
PROSITE	PS50975; ATP_GRASP; 1.
KW	Kinase.
SQ	SEQUENCE 319 AA; 36220 MW; D00E60DE4F288D1B CRC64;
Query Match	54.3%; Score 44; DB 2; Length 319;
Best Local Similarity	63.6%; Pred. No. 49;
Matches	7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy	3 LQEFLSHGGYV 13 : :
Dd	176 LQEFVNHGVI 186 : :
RESULT 34	
O9SBA5 ARATH	

Mon May 15 11:35:37 2006

```
KW Kinase.
SQ SEQUENCE 342 AA; 37312 MW; F24DEB3FEBEE6F3C CRC64;

Query Match 54.3%; Score 44; DB 2; Length 319;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13
| | | | | :
| | | | | :
Db 176 LOEFVNHGGVI 186

RESULT 35
Q7XBW0_ORYSA
ID Q7XBW0_ORYSA PRELIMINARY; PRT; 342 AA.
AC Q7XBW0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative inositol 1,3,4-trisphosphate 5/6-kinase.
GN ORFNames=OSJNBa0027L23.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wang R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017122; AAP55148.1; -; Genomic_DNA.
DR Gramene; Q7XBW0; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR008656; Insl134_P3_kin.
DR Pfam; PF05770; Insl134_P3_kin; 1.
KW Kinase.
SQ SEQUENCE 342 AA; 38072 MW; 3B99D2AB5B46FB93 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13
| | | | | :
| | | | | :
Db 185 LOEFVNHGGVI 195

RESULT 36
Q84Y01_MAIZE
ID Q84Y01_MAIZE PRELIMINARY; PRT; 342 AA.
AC Q84Y01;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol phosphate kinase.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22474773; PubMed=12586875; DOI=10.1104/pp.014258;
RA Shi J., Wang H., Wu Y., Hazebroek J., Meeley R.B., Ertl D.S.;
RT "The maize low-phytic acid mutant lpa2 is caused by mutation in an
RT inositol phosphate kinase gene.";
RL Plant Physiol. 131:507-515(2003).
DR EMBL; AY172635; AA017299.1; -; mRNA.
DR GO; GO:0016301; F:kinase activity; IEA.
```

```
KW Kinase.
SQ SEQUENCE 342 AA; 37312 MW; F24DEB3FEBEE6F3C CRC64;

Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13
| | | | | :
| | | | | :
Db 186 LOEFVNHGGVI 196

RESULT 37
Q8W3H7_ORYSA
ID Q8W3H7_ORYSA PRELIMINARY; PRT; 342 AA.
AC Q8W3H7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol 1,3,4-trisphosphate 5/6-kinase.
GN Name=OSJNBa0027L23.5;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Heiao J., Zismann T.R., Feidblyum T.V., Kalb E., Quackenbush J.,
RA VanAken S.E., Utterback T.R., Feldblyum C.M.;
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018929; AAL67584.1; -; Genomic_DNA.
DR Gramene; Q8W3H7; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR008656; Insl134_P3_kin.
DR Pfam; PF05770; Insl134_P3_kin; 1.
KW Kinase.
SQ SEQUENCE 342 AA; 38072 MW; 3B99D2AB5B46FB93 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 342;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LOEFLSHGGYV 13
| | | | | :
| | | | | :
Db 185 LOEFVNHGGVI 195

RESULT 38
Q92M44_RHIME
ID Q92M44_RHIME PRELIMINARY; PRT; 348 AA.
AC Q92M44;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PUTATIVE LOW SPECIFICITY L-THREONINE ALDOLASE PROTEIN
DE (EC 4.1.2.5).
GN OrderedLocusNames=R02812; ORFNames=SMC04029;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
```

```

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Maasu D.,
RA Pohl T., Portet-Leclerc D., Puehler A., Purnelle B., Ramepgerger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
DR EMBL; AL591791; CAC47391.1; -; Genomic DNA.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR001597; Beta_elim_lyase.
DR Pfam; PF01212; Beta_elim_lyase; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 348 AA; 37858 MW; 044516F6B5D5CA78 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 348;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CSLOEFLSHGGYVC 14
| | | | |
Db 90 CGAPEFFSHGARLC 103

RESULT 39
Q98E45 RHIL0 PRELIMINARY; PRT; 388 AA.
AC Q98E45;
AT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE M114416 protein.
OS OrderedLocusNames=m114416;
GN Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MAPF303099;
RC MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL Mesorhizobium loti";
RL DNA Res. 7:331-338 (2000).
DR EMBL; BA000012; BAB51075.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006076; Rad_oxred.
DR Pfam; PF01266; DAO; 1.
KW Complete proteome.
SQ SEQUENCE 388 AA; 41940 MW; 77B6F256B17490F5 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 388;
Best Local Similarity 72.7%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFLLSHGGY 12
| | | | |
Db 353 ALAEFLVHGGY 363

RESULT 40
Q95K70 MACFA PRELIMINARY; PRT; 393 AA.
AC Q95K70;
AT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Cerebellum cortex;
RC Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB066509; BAB62196.1; -; mRNA.
DR HSP; P08603; 1HFH.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS0923; SUSHI; 3.
KW Hypothetical protein.
SQ SEQUENCE 393 AA; 43314 MW; F964759AF2C75CF9 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 393;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CSLOEFLSHGGYVC 14
| | | | |
Db 146 CPLPPMVSHGDFVC 159

RESULT 41
Q6AJD7 DESPS PRELIMINARY; PRT; 413 AA.
AC Q6AJD7;
AT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Related to penicillin-binding protein 4 [Precursor].
GN OrderedLocusNames=DP2814;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=LSV54 / DSM 12343;
RC PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:987-902 (2004).
DR EMBL; CR522870; CAG37543.1; -; Genomic DNA.
DR GO; GO:0004185; F:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000667; Peptidase_S13.
DR Pfam; PF02113; Peptidase_S13; 1.
DR PRINTS; PR00922; DADACBPTASE3.
DR TIGRFAMs; TIGR00666; BPP4; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 46188 MW; B88E4A9149B52902 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 413;
Best Local Similarity 72.7%; Pred. No. 65;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFLLSHGGY 12

```

```
Db 34 SLQEDIDGGY 44
||||| : |||||
||||| : |||||

RESULT 42
MYCS RAT STANDARD; PRT; 429 AA.
AC P23999;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE S-myc protein.
GN Name=MyCS;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90083224; PubMed=2594755;
RA Sugiyama A., Kume A., Nemoto K., Lee S.Y., Asami Y., Nemoto F.,
RA Nishimura S., Kuchino Y.;
RT "Isolation and characterization of s-myc, a member of the rat myc gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9144-9148(1989).
CC -!- FUNCTION: Has apoptosis-inducing activity.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; M29069; AAA1645.1; -; Genomic_DNA.
DR PIR; A36220; A36220.
DR HSSP; P25912; 1HLQ.
DR TRANSFAC; T03550; -.
DR Ensembl; ENSRNOG0000003085; Rattus norvegicus.
DR RGD; 3133; Mycs.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR002418; TF_Myc.
DR InterPro; IPR012682; TF_Myc_N.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF01056; MYC_N; 1.
DR PRINTS; PR00044; LEUZIPRMYC.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS08888; HLH; 1.
KW Apoptosis; DNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 360 399 Helix-loop-helix motif.
FT DOMAIN 398 419 Leucine-zipper (Potential).
FT DNA_BIND 347 359 Basic motif.
FT MOD_RES 36 36 Phosphotyrosine (by Tyr-kinases) (By
FT similarity).
SQ SEQUENCE 429 AA; 47002 MW; 1E79FF90747FD0FA CRC64;

Query Match 54.3%; Score 44; DB 1; Length 429;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYV 13
|||:|||||
Db 227 SLEDFLSNGYV 238

RESULT 43
Q7N4F2 PHOLL
ID Q7N4F2_PHOLL PRELIMINARY; PRT; 430 AA.
AC Q7N4F2;

DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Similar to unknown protein.
GN OrderedLocustNames=plu2387;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN 1
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Tauchard E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Dourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571867; CAE14715.1; -; Genomic_DNA.
DR Photolast; plu2387; -
DR InterPro; IPR007067; Phage_sheath_1.
DR Pfam; PF04984; Phage_sheath_1; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 46718 MW; 45A307E1169B400F CRC64;

Query Match 54.3%; Score 44; DB 2; Length 430;
Best Local Similarity 53.8%; Pred. No. 68;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SLQEFSLSHGGYV 14
:||:|||||
Db 99 AVQHYFSHGGGV 111

RESULT 44
QSRM2 PONPY
ID QSRM2_PONPY PRELIMINARY; PRT; 489 AA.
AC QSRM2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469E106.
GN Name=DKFZp469E106;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN 1
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RG The German cDNA Consortium;
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Pobo G., Han M., Wiemann S.;
RA Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR859729; CAH91888.1; -; mRNA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 4.
DR SMART; SM00032; CCP; 4.
DR PROSITE; PS00923; SUSHI; 4.
KW Hypothetical protein.
SQ SEQUENCE 489 AA; 53697 MW; 5241CA08BFEE29990 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 489;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYV 14
|||:|||||
Db 242 CPLPPMVSHGDFVC 255
```

```
DR EMBL; AL359979; CAH73991.1; JOINED; Genomic_DNA.
DR EMBL; AL359733; CAI12801.1; JOINED; Genomic_DNA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 4.
DR SMART; SM00032; CCP; 4.
DR PROSITE; PS0923; SUSHI; 4.
SQ SEQUENCE 490 AA; 53777 MW; 15392C2774BEFF22 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 490;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
| | : | | |
DB 243 CPLPPMVSHGDFVC 256

RESULT 47
Q4L8P7_STAHP PRELIMINARY; PRT; 701 AA.
AC Q4L8P7_STAHP PRELIMINARY;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Similar to transcription regulatory protein.
GN ORFNames=SH0669;
OS Staphylococcus haemolyticus (strain JCS1435).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=279808;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JCS1435;
RA Takeuchi F., Matanabe S., Baba T., Yuzawa H., Ito T., Cui L.,
RA Morimoto Y., Kuroda M., Takahashi M., Anai A., Baba S., Fukui S.,
RA Lee J.C., Hiramoto K.;
RT "Whole genome sequencing of Staphylococcus haemolyticus uncovers
RT extreme plasticity of its genome and dynamism in the evolution of
RT human-colonizing staphylococcal species."
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006716; BA503978.1; -; Genomic_DNA.
SQ SEQUENCE 701 AA; 82194 MW; D2C488B066767F05 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 701;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
| | : | | |
DB 533 CHLYQLMRHGGYVC 546

RESULT 48
O07363_CHLTR PRELIMINARY; PRT; 123 AA.
AC O07363;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L2/434/Bu;
RA Raggiaschi R., Ratti G.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U74759; AAB58903.1; -; Genomic_DNA.
DR HSSP; P05055; ISRO.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.

Q6UX62_HUMAN PRELIMINARY; PRT; 490 AA.
AC Q6UX62;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE YHGM196.
GN ORFNames=UNQ196;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brueh J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Sehagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Matanabe C., Wisand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment."
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY358495; AAQ88859.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00084; Sushi; 4.
DR SMART; SM00032; CCP; 4.
DR PROSITE; PS0923; SUSHI; 4.
SQ SEQUENCE 490 AA; 53919 MW; 950254448E4267D1 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 490;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSLQEFSLSHGGYVC 14
| | : | | |
DB 245 CPLPPMVSHGDFVC 258

RESULT 46
Q5VX71_HUMAN PRELIMINARY; PRT; 490 AA.
AC Q5VX71;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Novel protein (FLJ10052).
GN Novel=RP11-239E10.4; ORFNames=RP11-239E10.4-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Pelan S.;
RA Pelan S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Almeida J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359733; CAH73991.1; -; Genomic_DNA.
DR EMBL; AL359979; CAI12801.1; -; Genomic_DNA.
```

```

DR InterPro; IPR000110; Ribosomal_S1.
DR InterPro; IPR003029; S1.
DR Pfam; PF00575; S1; 1.
DR PRINTS; PR00861; RIBOSOMALS1.
DR SMART; SM00316; S1; 1.
DR PROSITE; PS0126; S1; 1.
DR KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 123 AA; 13376 MW; 993084CF4DDA319C CRC64;

Query Match 53.1%; Score 43; DB 2; Length 123;
Best Local Similarity 70.0%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGG 11
Db 100 SIKEFLAHGG 109

RESULT 49
Q4RU12_TETNG PRELIMINARY; PRT; 219 AA.
ID Q4RU12;
AC Q4RU12;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 12 SCAR1496, whole genome shotgun sequence.
GN ORFNames=GSTENG0028997001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphia; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouteau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014996; CAG08120.1; -; Genomic DNA.
SQ SEQUENCE 219 AA; 24161 MW; 7F62B3CBC968EC01 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 219;
Best Local Similarity 72.7%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SLQEFLSHGGY 12
Db 167 SAQEFLLTHEGY 177

RESULT 50
Q759D9_ASHGO PRELIMINARY; PRT; 469 AA.
ID Q759D9_ASHGO

```

```

AC Q759D9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ADR338CP
GN Name=ADR338C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016817; AAS52258.1; -; Genomic_DNA.
DR AGD; ADR338C; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR004595; Ssl1.
DR InterPro; IPR007198; Ssl1-like.
DR InterPro; IPR012170; TFIIH_SSL1.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04056; Ssl1; 1.
DR PIRSF; PIRSF015919; TFIIH_SSL1; 1.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMs; TIGR00622; ssl1; 1.
DR PROSITE; PS50234; VWF_A; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR Complete proteome.
SQ SEQUENCE 469 AA; 52811 MW; 44638920D54DDF54 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 469;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CSLQEFLSHGGYVC 14
Db 345 CTCHSRLVHGGYFC 358

Search completed: May 12, 2006, 10:52:31
Job time : 130.077 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 72.6496 Seconds
(without alignments)

60.479 Million cell updates/sec

Title: US-10-632-388-62

Perfect score: 56

Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A_Geneseq_21.*

1: _geneseqp1980s.*

2: _geneseqp1990s.*

3: _geneseqp2000s.*

4: _geneseqp2001s.*

5: _geneseqp2002s.*

6: _geneseqp2003as.*

7: _geneseqp2003bs.*

8: _geneseqp2004s.*

9: _geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	10	3	AAb17006 TPO-mimet
2	56	100.0	10	5	ABb72892 TPO mimet
3	56	100.0	10	7	ADJ73043 TPO mimet
4	56	100.0	10	8	ADJ52678 CHI delet
5	56	100.0	10	8	ADJ51639 CHI delet
6	56	100.0	12	2	AAW36787 Thrombopo
7	56	100.0	12	2	AAW36781 Thrombopo
8	56	100.0	12	4	AAU26014 Human thr
9	56	100.0	13	2	AAW36779 Thrombopo
10	56	100.0	13	2	AAW36779 Human thr
11	56	100.0	13	4	AAU26008 Human thr
12	56	100.0	13	4	AAU26012 Human thr
13	56	100.0	13	4	AAU26015 Human thr
14	56	100.0	13	7	ADJ73003 TPO mimet
15	56	100.0	13	8	ADJ52638 CHI delet
16	56	100.0	13	8	ADJ51599 CHI delet
17	56	100.0	14	2	AAW09463 Thrombopo
18	56	100.0	14	2	AAW09468 Thrombopo
19	56	100.0	14	2	AAW33030 Thrombopo
20	56	100.0	14	2	AAW33034 Thrombopo
21	56	100.0	14	2	AAW36782 Thrombopo
22	56	100.0	14	2	AAW36774 Thrombopo
23	56	100.0	14	2	AAW36788 Thrombopo
24	56	100.0	14	2	ADI24843 AF 12505

25	56	100.0	14	3	AAy96515	Thrombopo
26	56	100.0	14	3	AAb16962	TPO-mimet
27	56	100.0	14	3	AAb16968	TPO-mimet
28	56	100.0	14	4	AAU26009	Human thr
29	56	100.0	14	4	AAU26013	Human thr
30	56	100.0	14	4	AAU26006	Human thr
31	56	100.0	14	4	AAU25827	Human thr
32	56	100.0	14	4	AAU26010	Human thr
33	56	100.0	14	4	AAU26004	Human thr
34	56	100.0	14	5	ABb72854	TPO mimet
35	56	100.0	14	5	ABb72853	TPO mimet
36	56	100.0	14	5	ABb72853	Thrombopo
37	56	100.0	14	5	AAE18011	Human lig
38	56	100.0	14	6	ABR62907	TPO recep
39	56	100.0	14	7	ABR62907	Thrombopo
40	56	100.0	14	7	ADc33697	Erythro
41	56	100.0	14	7	ADJ73004	TPO mimet
42	56	100.0	14	7	ADN59652	Thrombopo
43	56	100.0	14	8	ADJ52639	CHI delet
44	56	100.0	14	8	ADJ51600	CHI delet
45	56	100.0	14	8	ADL27293	Amino aci
46	56	100.0	14	8	ADMT2503	TPO mimet
47	56	100.0	14	8	ADMT2483	TPO mimet
48	56	100.0	14	8	ADMT2495	TPO mimet
49	56	100.0	14	8	ADMT2487	TPO mimet
50	56	100.0	14	8	ADMT2497	TPO mimet
51	56	100.0	14	8	ADQ16584	Agonist T
52	56	100.0	14	8	ADT92482	Thrombopo
53	56	100.0	14	9	ADU70210	Thrombopo
54	56	100.0	14	9	ADU70206	Thrombopo
55	56	100.0	14	9	ADU75982	Peptide-b
56	56	100.0	14	9	ADU75979	Thrombopo
57	56	100.0	14	9	ADV44319	Agonist T
58	56	100.0	14	9	AEb12792	TPO mimet
59	56	100.0	15	2	AAW35416	Thrombopo
60	56	100.0	15	2	AAW36784	Thrombopo
61	56	100.0	15	2	AAW36780	Thrombopo
62	56	100.0	15	2	AAW36776	Thrombopo
63	56	100.0	15	2	AAW66717	Thrombopo
64	56	100.0	15	2	AAW66718	Peptide c
65	56	100.0	15	2	AAW66714	Peptide c
66	56	100.0	15	2	AAW66721	Peptide c
67	56	100.0	15	2	AAW66712	Peptide c
68	56	100.0	15	3	AAAB20684	Thrombocy
69	56	100.0	15	4	AAU26022	Human thr
70	56	100.0	15	4	AAU25996	Human thr
71	56	100.0	15	4	AAU26026	Human thr
72	56	100.0	15	4	AAU26011	Human thr
73	56	100.0	15	4	AAU26020	Human thr
74	56	100.0	15	4	AAU25831	Human thr
75	56	100.0	15	4	AAU26023	Human thr
76	56	100.0	15	4	AAU26007	Human thr
77	56	100.0	15	5	ABP51670	Thrombopo
78	56	100.0	15	7	ABR62908	Thrombopo
79	56	100.0	15	8	ADMT2485	TPO mimet
80	56	100.0	15	8	ADMT2479	TPO mimet
81	56	100.0	15	8	ADMT2502	TPO mimet
82	56	100.0	15	8	ADMT2492	TPO mimet
83	56	100.0	15	8	ADMT2478	TPO mimet
84	56	100.0	15	8	ADMT2533	TPO mimet
85	56	100.0	15	8	ADMT2496	TPO mimet
86	56	100.0	15	8	ADMT2490	TPO mimet
87	56	100.0	15	8	ADMT2486	TPO mimet
88	56	100.0	15	8	ADMT2491	TPO mimet
89	56	100.0	15	8	ADMT2522	TPO mimet
90	56	100.0	15	8	ADMT2523	TPO mimet
91	56	100.0	15	8	ADMT2493	TPO mimet
92	56	100.0	15	8	ADMT2482	TPO mimet
93	56	100.0	15	8	ADMT2494	TPO mimet
94	56	100.0	15	8	ADMT2494	TPO mimet
95	56	100.0	15	8	ADQ16585	TPO mimet
96	56	100.0	15	8	ADT92483	Modified
97	56	100.0	15	9	ADU70209	Thrombopo
98	56	100.0	15	9	ADU75981	Peptide-b

98	56	100.0	15	9	ADV44320	Agv44320 Agonist T	171	56	100.0	18	9	ABE12939	Abel2939 TPO mimet
99	56	100.0	15	9	ABE12793	Abel12793 TPO mimet	172	56	100.0	18	9	ABE12832	Abel2832 TPO mimet
100	56	100.0	16	2	AAW19534	AAW19534 Thrombopo	173	56	100.0	18	9	ABE12852	Abel2852 TPO mimet
101	56	100.0	16	2	AAW33035	AAW33035 Thrombopo	174	56	100.0	18	9	ABE12816	Abel2816 TPO mimet
102	56	100.0	16	2	AAW36775	AAW36775 Thrombopo	175	56	100.0	18	9	ABE12818	Abel2818 TPO mimet
103	56	100.0	16	2	AAW36771	AAW36771 Thrombopo	176	56	100.0	18	9	ABE12828	Abel2828 TPO mimet
104	56	100.0	16	2	AAW66709	AAW66709 Peptide c	177	56	100.0	18	9	ABE12836	Abel2836 TPO mimet
105	56	100.0	16	2	AAW66713	AAW66713 Peptide c	178	56	100.0	18	9	ABE12822	Abel2822 TPO mimet
106	56	100.0	16	2	AAW66733	AAW66733 Peptide c	179	56	100.0	18	9	ABE12844	Abel2844 TPO mimet
107	56	100.0	16	2	AAW66716	AAW66716 Peptide c	180	56	100.0	18	9	ABE12838	Abel2838 TPO mimet
108	56	100.0	16	4	AAU26021	AAU26021 Human thr	181	56	100.0	19	5	ABW73391	Abw73391 TPO-mimet
109	56	100.0	16	4	AAU26005	AAU26005 Human thr	182	56	100.0	19	5	ABW73390	Abw73390 TPO-mimet
110	56	100.0	16	4	AAU26043	AAU26043 Human thr	183	56	100.0	20	3	ABW18003	Abw18003 Pe-TMP pe
111	56	100.0	16	4	AAU25832	AAU25832 Human thr	184	56	100.0	20	3	ABW17929	Abw17929 TPO-mimet
112	56	100.0	16	8	ADM72532	ADM72532 TPO mimet	185	56	100.0	20	3	ABW73403	Abw73403 TPO mimet
113	56	100.0	16	8	ADM72484	ADM72484 TPO mimet	186	56	100.0	21	7	ADN59687	Adn59687 Thrombopo
114	56	100.0	18	3	ABW16957	ABW16957 PEGylated	187	56	100.0	22	7	ADN59830	Adn59830 TWP pepti
115	56	100.0	18	3	ABW16956	ABW16956 PEGylated	188	56	100.0	22	7	ADN59819	Adn59819 TWP pepti
116	56	100.0	18	5	ABW1687	ABW1687 TPO mimet	189	56	100.0	22	8	ADQ16714	Adq16714 Immunoglo
117	56	100.0	18	5	ABW1689	ABW1689 TPO mimet	190	56	100.0	22	8	ADQ16713	Adq16713 Immunoglo
118	56	100.0	18	5	ABW1688	ABW1688 TPO mimet	191	56	100.0	22	8	ADQ16709	Adq16709 Immunoglo
119	56	100.0	18	5	ABW1677	ABW1677 TPO mimet	192	56	100.0	22	8	ADQ16706	Adq16706 Immunoglo
120	56	100.0	18	5	ABW1686	ABW1686 TPO mimet	193	56	100.0	22	8	ADQ16699	Adq16699 TPO mimet
121	56	100.0	18	5	ABW1674	ABW1674 TPO mimet	194	56	100.0	22	8	ADQ16712	Adq16712 Immunoglo
122	56	100.0	18	5	ABW1693	ABW1693 TPO mimet	195	56	100.0	22	8	ADQ16707	Adq16707 Immunoglo
123	56	100.0	18	5	ABW1684	ABW1684 TPO mimet	196	56	100.0	22	8	ADQ16711	Adq16711 Immunoglo
124	56	100.0	18	5	ABW1683	ABW1683 TPO mimet	197	56	100.0	22	8	ADQ16708	Adq16708 Immunoglo
125	56	100.0	18	5	ABW1685	ABW1685 TPO mimet	198	56	100.0	22	8	ADQ16710	Adq16710 Immunoglo
126	56	100.0	18	5	ABW1691	ABW1691 TPO mimet	199	56	100.0	22	9	ADV44435	Adv44435 Modified
127	56	100.0	18	5	ABW1673	ABW1673 TPO mimet	200	56	100.0	22	9	ADV44449	Adv44449 Anti-teta
128	56	100.0	18	5	ABW1690	ABW1690 TPO mimet	201	56	100.0	22	9	ADV44443	Adv44443 Anti-teta
129	56	100.0	18	5	ABW1675	ABW1675 TPO mimet	202	56	100.0	22	9	ADV44444	Adv44444 Anti-teta
130	56	100.0	18	5	ABW1692	ABW1692 TPO mimet	203	56	100.0	22	9	ADV44448	Adv44448 Anti-teta
131	56	100.0	18	7	ADN59663	Adn59663 Thrombopo	204	56	100.0	22	9	ADV44442	Adv44442 Anti-teta
132	56	100.0	18	7	ADN59812	Adn59812 Thrombopo	205	56	100.0	22	9	ADV44447	Adv44447 Anti-teta
133	56	100.0	18	8	ADQ16611	Adq16611 TPO mimet	206	56	100.0	22	9	ADV44445	Adv44445 Anti-teta
134	56	100.0	18	8	ADQ16619	Adq16619 TPO mimet	207	56	100.0	22	9	ADV44450	Adv44450 Anti-teta
135	56	100.0	18	8	ADQ16621	Adq16621 TPO mimet	208	56	100.0	22	9	ADV44446	Adv44446 Anti-teta
136	56	100.0	18	8	ADQ16641	Adq16641 TPO mimet	209	56	100.0	22	9	ABE12918	Abel2918 TPO mimet
137	56	100.0	18	8	ADQ16646	Adq16646 TPO mimet	210	56	100.0	22	9	ABE12917	Abel2917 TPO mimet
138	56	100.0	18	8	ADQ16607	Adq16607 TPO mimet	211	56	100.0	22	9	ABE12919	Abel2919 TPO mimet
139	56	100.0	18	8	ADQ16615	Adq16615 TPO mimet	212	56	100.0	22	9	ABE12986	Abel2986 TT antibo
140	56	100.0	18	8	ADQ16627	Adq16627 TPO mimet	213	56	100.0	22	9	ABE12920	Abel2920 TPO mimet
141	56	100.0	18	8	ADQ16625	Adq16625 TPO mimet	214	56	100.0	22	9	ABE12916	Abel2916 TPO mimet
142	56	100.0	18	8	ADQ16617	Adq16617 TPO mimet	215	56	100.0	22	9	ABE12914	Abel2914 TPO mimet
143	56	100.0	18	8	ADQ16629	Adq16629 TPO mimet	216	56	100.0	22	9	ABE12921	Abel2921 TPO mimet
144	56	100.0	18	8	ADQ16613	Adq16613 TPO mimet	217	56	100.0	22	9	ABE12907	Abel2907 TT antibo
145	56	100.0	18	8	ADQ16623	Adq16623 TPO mimet	218	56	100.0	22	9	ABE12915	Abel2915 TPO mimet
146	56	100.0	18	8	ADQ16605	Adq16605 TPO mimet	219	56	100.0	25	7	ADN59708	Adn59708 Thrombopo
147	56	100.0	18	8	ADQ16609	Adq16609 TPO mimet	220	56	100.0	28	3	ABW17285	Abw17285 TPO-mimet
148	56	100.0	18	9	ADV44355	Adv44355 Agonist T	221	56	100.0	28	5	ABW1682	Abw1682 TPO mimet
149	56	100.0	18	9	ADV44345	Adv44345 Agonist T	222	56	100.0	28	7	ADJ73013	Adj73013 TPO mimet
150	56	100.0	18	9	ADV44351	Adv44351 Agonist T	223	56	100.0	28	8	ADJ52648	Adj52648 CH1 delet
151	56	100.0	18	9	ADV44357	Adv44357 Agonist T	224	56	100.0	28	8	ADJ51609	Adj51609 CH1 delet
152	56	100.0	18	9	ADV44349	Adv44349 Agonist T	225	56	100.0	28	8	ADQ16636	Adq16636 Tetanus t
153	56	100.0	18	9	ADV44343	Adv44343 Agonist T	226	56	100.0	28	9	ADV44374	Adv44374 Modified
154	56	100.0	18	9	ADV44347	Adv44347 Agonist T	227	56	100.0	28	9	ABE12847	Abel2847 Antibody
155	56	100.0	18	9	ADV44384	Adv44384 Agonist T	228	56	100.0	29	3	ABW16971	Abw16971 TPO-mimet
156	56	100.0	18	9	ADV44353	Adv44353 Agonist T	229	56	100.0	29	3	ABW16975	Abw16975 TPO-mimet
157	56	100.0	18	9	ADV44363	Adv44363 Agonist T	230	56	100.0	29	3	ABW16976	Abw16976 TPO-mimet
158	56	100.0	18	9	ADV44361	Adv44361 Agonist T	231	56	100.0	29	3	ABW17286	Abw17286 TPO-mimet
159	56	100.0	18	9	ADV44379	Adv44379 Agonist T	232	56	100.0	29	3	ABW16970	Abw16970 TPO-mimet
160	56	100.0	18	9	ADV44344	Adv44344 Agonist T	233	56	100.0	29	5	ABW72862	Abw72862 TPO mimet
161	56	100.0	18	9	ADV44359	Adv44359 Agonist T	234	56	100.0	29	5	ABW72857	Abw72857 TPO mimet
162	56	100.0	18	9	ADV44365	Adv44365 Agonist T	235	56	100.0	29	5	ABW72861	Abw72861 TPO mimet
163	56	100.0	18	9	ADV44367	Adv44367 Agonist T	236	56	100.0	29	5	ABW72856	Abw72856 TPO mimet
164	56	100.0	18	9	ADV44468	Adv44468 Agonist T	237	56	100.0	29	7	ADJ73011	Adj73011 TPO mimet
165	56	100.0	18	9	ABE12826	Abel2826 TPO mimet	238	56	100.0	29	7	ADJ73007	Adj73007 TPO mimet
166	56	100.0	18	9	ABE12820	Abel2820 TPO mimet	239	56	100.0	29	7	ADJ73006	Adj73006 TPO mimet
167	56	100.0	18	9	ABE12840	Abel2840 TPO mimet	240	56	100.0	29	8	ADJ52642	Adj52642 CH1 delet
168	56	100.0	18	9	ABE12857	Abel2857 Antibody	241	56	100.0	29	8	ADJ52646	Adj52646 CH1 delet
169	56	100.0	18	9	ABE12834	Abel2834 TPO mimet	242	56	100.0	29	8	ADJ52641	Adj52641 CH1 delet
170	56	100.0	18	9	ABE12830	Abel2830 TPO mimet	243	56	100.0	29	8	ADJ51603	Adj51603 CH1 delet

244	56	100.0	29	8	ADJ51602	Adj51602	CH1 delet	317	56	100.0	234	9	AEb12912	Antibody
245	56	100.0	29	8	ADJ51607	Adj51607	CH1 delet	318	56	100.0	247	3	ABb16958	FC-TMP pr
246	56	100.0	30	3	ABb17287	ABb17287	TPO-mimet	319	56	100.0	247	3	ABb16958	TMP-FC pr
247	56	100.0	31	3	ABb17288	ABb17288	TPO-mimet	320	56	100.0	247	5	ABb73411	FC-TPO m1
248	56	100.0	31	3	ABb16974	ABb16974	TPO-mimet	321	56	100.0	247	5	ABb73411	TMP-FC am
249	56	100.0	31	3	ABb16973	ABb16973	TPO-mimet	322	56	100.0	249	3	ADV44440	pAX116 va
250	56	100.0	31	5	ABb72860	ABb72860	TPO mimet	323	56	100.0	268	3	ABb16959	FC-TMP-TM
251	56	100.0	31	5	ABb72859	ABb72859	TPO mimet	324	56	100.0	268	3	ABb73412	FC-TMP-TM
252	56	100.0	31	7	ADJ73009	Adj73009	TPO mimet	325	56	100.0	269	3	AAy96531	Human IGG
253	56	100.0	31	7	ADJ73010	Adj73010	TPO mimet	326	56	100.0	269	3	ABb16960	TMP-TMP-F
254	56	100.0	31	8	ADJ52644	Adj52644	CH1 delet	327	56	100.0	269	5	ABb73413	TMP-TMP-F
255	56	100.0	31	8	ADJ52645	Adj52645	CH1 delet	328	56	100.0	282	9	AEb12930	Antibody
256	56	100.0	31	8	ADJ51606	Adj51606	CH1 delet	329	56	100.0	425	9	ADV44459	Anti-teta
257	56	100.0	31	8	ADJ51605	Adj51605	CH1 delet	330	56	100.0	472	5	ABP51695	SG1.1-TPO
258	56	100.0	32	3	AAy96520	AAy96520	Thrombopo	331	56	100.0	472	8	ADQ16647	Immunoglob
259	56	100.0	32	3	ABb17289	ABb17289	TPO-mimet	332	56	100.0	472	9	ADV44385	SG1.1 hea
260	56	100.0	32	3	ABb17297	ABb17297	TPO-mimet	333	56	100.0	472	9	AEb12858	Antibody
261	56	100.0	33	3	AAb17290	AAb17290	TPO-mimet	334	53	94.6	14	3	ABb16969	TPO-mimet
262	56	100.0	33	3	AAy96527	AAy96527	Thrombopo	335	53	94.6	14	5	ABb72855	TPO mimet
263	56	100.0	34	3	AAb17291	AAb17291	TPO-mimet	336	53	94.6	14	7	ADJ73005	TPO mimet
264	56	100.0	35	3	AAb17292	AAb17292	TPO-mimet	337	53	94.6	14	8	ADJ52640	CH1 delet
265	56	100.0	35	3	AAy96525	AAy96525	Thrombopo	338	53	94.6	14	8	ADJ51601	CH1 delet
266	56	100.0	36	3	AAy96523	AAy96523	Thrombopo	339	53	94.6	15	2	AAW66719	Peptide c
267	56	100.0	36	3	AAy96524	AAy96524	Thrombopo	340	52	92.9	13	4	AAU26018	Human thr
268	56	100.0	36	3	AAy96526	AAy96526	Thrombopo	341	52	92.9	13	4	AAU26035	Human thr
269	56	100.0	36	3	ABb17307	ABb17307	TPO-mimet	342	52	92.9	13	8	ADm72525	TPO mimet
270	56	100.0	36	3	ABb17293	ABb17293	TPO-mimet	343	52	92.9	13	8	ADm72488	TPO mimet
271	56	100.0	36	3	ABb17303	ABb17303	TPO-mimet	344	52	92.9	13	8	ADm72489	TPO mimet
272	56	100.0	36	3	ABb16963	ABb16963	TPO-mimet	345	52	92.9	14	2	AAW66773	Thrombopo
273	56	100.0	36	3	AAb17301	AAb17301	TPO-mimet	346	52	92.9	14	2	AAW66715	Peptide c
274	56	100.0	36	3	ABb17306	ABb17306	TPO-mimet	347	52	92.9	14	2	AAW66730	Peptide c
275	56	100.0	36	5	ABb72403	ABb72403	TPO-mimet	348	52	92.9	14	2	AAW66730	Peptide c
276	56	100.0	37	3	ABb17294	ABb17294	TPO-mimet	349	52	92.9	14	4	AAU26019	Human thr
277	56	100.0	38	3	ABb17295	ABb17295	TPO-mimet	350	52	92.9	14	4	AAU26036	Human thr
278	56	100.0	39	3	ABb17304	ABb17304	TPO-mimet	351	52	92.9	14	4	AAU26037	Human thr
279	56	100.0	39	3	ABb17305	ABb17305	TPO-mimet	352	52	92.9	14	8	ADm72526	TPO mimet
280	56	100.0	40	3	ABb17302	ABb17302	TPO-mimet	353	52	92.9	15	2	AAW66731	Peptide c
281	56	100.0	41	5	ABb73389	ABb73389	TPO-mimet	354	52	92.9	15	4	AAU26038	Human thr
282	56	100.0	41	5	ABb73388	ABb73388	TPO-mimet	355	52	92.9	18	7	ADN59815	Thrombopo
283	56	100.0	42	3	AAy96530	AAy96530	Thrombopo	356	52	92.9	18	7	ADN59860	Thrombopo
284	56	100.0	42	3	AAy96530	AAy96530	Thrombopo	357	52	92.9	18	7	ADN59860	Thrombopo
285	56	100.0	42	3	ABb17296	ABb17296	Synthetic	358	52	92.9	18	7	ADN59860	Thrombopo
286	56	100.0	42	3	ABb17308	ABb17308	TPO-mimet	359	52	92.9	18	7	ADN59860	Thrombopo
287	56	100.0	42	3	ABb17282	ABb17282	TPO-mimet	360	52	92.9	19	2	AAW09494	Thrombopo
288	56	100.0	42	3	ABb17281	ABb17281	TPO-mimet	361	52	92.9	19	2	AAW09494	Thrombopo
289	56	100.0	43	5	ABb73404	ABb73404	TMP-TMP g	362	52	92.9	19	2	AAW09494	Thrombopo
290	56	100.0	43	5	ABb73404	ABb73404	TMP-TMP g	363	52	92.9	19	2	AAW09494	Thrombopo
291	56	100.0	60	3	ABb73405	ABb73405	Anti-teta	364	52	92.9	19	2	AAW09494	Thrombopo
292	56	100.0	60	3	ABb73405	ABb73405	Anti-teta	365	52	92.9	19	2	AAW09494	Thrombopo
293	56	100.0	122	9	ADV44474	ADV44474	Anti-teta	366	52	92.9	19	2	AAW09494	Thrombopo
294	56	100.0	122	9	AEb12946	AEb12946	Antibody	367	52	92.9	19	4	AAU25864	Human thr
295	56	100.0	128	8	ADQ16705	ADQ16705	Modified	368	52	92.9	19	4	AAU25864	Human thr
296	56	100.0	128	8	ADQ16705	ADQ16705	Modified	369	52	92.9	19	4	AAU25864	Human thr
297	56	100.0	128	9	ADV44466	ADV44466	Anti-teta	370	52	92.9	19	4	AAU25864	Human thr
298	56	100.0	128	9	ADV44467	ADV44467	Anti-teta	371	52	92.9	22	7	ADN59827	TMP pepi
299	56	100.0	128	9	ADV44467	ADV44467	Anti-teta	372	52	92.9	22	7	ADN59827	TMP pepi
300	56	100.0	128	9	ADV44467	ADV44467	Anti-teta	373	52	92.9	22	7	ADN59827	TMP pepi
301	56	100.0	128	9	AEb12934	AEb12934	Antibody	374	52	92.9	25	7	ADN59722	Thrombopo
302	56	100.0	128	9	AEb12934	AEb12934	Antibody	375	52	92.9	25	7	ADN59722	Thrombopo
303	56	100.0	128	9	AEb12933	AEb12933	Antibody	376	52	92.9	25	7	ADN59722	Thrombopo
304	56	100.0	128	9	AEb12933	AEb12933	Antibody	377	52	92.9	25	7	ADN59722	Thrombopo
305	56	100.0	128	9	AEb12936	AEb12936	Antibody	378	51	91.1	18	7	ADN59751	Peptide-v
306	56	100.0	128	9	AEb12937	AEb12937	Antibody	379	51	91.1	18	7	ADN59751	Peptide-v
307	56	100.0	128	9	AEb12938	AEb12938	Antibody	380	51	91.1	19	2	AAW09493	Thrombopo
308	56	100.0	129	6	ABG71751	ABG71751	Antibody	381	51	91.1	19	2	AAW09493	Thrombopo
309	56	100.0	131	6	ABG71753	ABG71753	Antibody	382	51	91.1	19	2	AAW09493	Thrombopo
310	56	100.0	132	9	ADV44473	ADV44473	Anti-teta	383	51	91.1	19	4	AAU25861	Human thr
311	56	100.0	133	6	ABG71752	ABG71752	Antibody	384	51	91.1	19	4	AAU25861	Human thr
312	56	100.0	135	6	ABG71749	ABG71749	Antibody	385	51	91.1	19	4	AAU25861	Human thr
313	56	100.0	143	6	ABG71750	ABG71750	Antibody	386	51	91.1	19	4	AAU25861	Human thr
314	56	100.0	144	6	ABG71748	ABG71748	Antibody	387	51	91.1	25	7	ADN59736	Thrombopo
315	56	100.0	150	9	AEb12945	AEb12945	Antibody	388	50	89.3	10	2	AAW09469	Thrombopo
316	56	100.0	225	8	ADQ16704	ADQ16704	Modified	389	50	89.3	10	2	AAW09469	Thrombopo

390	50	89.3	10	2	AAW36772	Thrombopo	463	49	87.5	19	4	AAU25870	Human thr
391	50	89.3	10	3	AAU17005	TPO-mimet	464	49	87.5	19	4	AAU25821	Human thr
392	50	89.3	10	4	AAU25839	Human thr	465	49	87.5	19	5	ABB72907	TPO mimet
393	50	89.3	10	5	AAU25840	Human thr	466	49	87.5	19	7	ADJ73059	TPO mimet
394	50	89.3	10	7	ABB72891	TPO mimet	467	49	87.5	19	8	ADJ52694	CH1 delet
395	50	89.3	10	7	ADJ73042	TPO mimet	468	49	87.5	19	8	ADJ51655	CH1 delet
396	50	89.3	10	8	ADJ52677	CH1 delet	469	49	87.5	22	7	ADN59826	TMP pepti
397	50	89.3	10	8	ADJ51638	CH1 delet	470	49	87.5	22	7	ADN59836	TMP pepti
398	50	89.3	14	8	ADJ51637	TPO mimet	471	49	87.5	22	7	ADN59836	TMP pepti
399	50	89.3	15	8	ADM72499	TPO mimet	472	49	87.5	22	7	ADN59821	TMP pepti
400	50	89.3	15	8	AAU26024	Human thr	473	49	87.5	22	7	ADN59838	TMP pepti
401	50	89.3	15	8	ADM72498	TPO mimet	474	49	87.5	22	7	ADN59839	TMP pepti
402	50	89.3	18	7	ADN59664	Thrombopo	475	49	87.5	22	7	ADN59834	TMP pepti
403	50	89.3	18	7	ADN59668	Thrombopo	476	49	87.5	23	7	ADN59795	Peptide-v
404	50	89.3	18	7	ADN59666	Thrombopo	477	49	87.5	23	7	ADN59777	Peptide-v
405	50	89.3	22	7	ADN59835	TMP pepti	478	49	87.5	25	7	ADN59742	Thrombopo
406	50	89.3	22	7	ADN59831	TMP pepti	479	49	87.5	25	7	ADN59716	Thrombopo
407	50	89.3	22	7	ADN59833	TMP pepti	480	49	87.5	25	7	ADN59724	Thrombopo
408	50	89.3	25	7	ADN59710	Thrombopo	481	49	87.5	25	7	ADN59691	Thrombopo
409	50	89.3	25	7	ADN59718	Thrombopo	482	49	87.5	25	7	ADN59700	Thrombopo
410	50	89.3	25	7	ADN59714	Thrombopo	483	49	87.5	25	7	ADN59720	Thrombopo
411	49	87.5	10	2	AAW09472	Thrombopo	484	49	87.5	25	7	ADN59744	Thrombopo
412	49	87.5	10	4	AAW36623	Thrombopo	485	49	87.5	36	7	ADN59765	Peptide-v
413	49	87.5	13	3	AAU25842	Human thr	486	49	87.5	41	7	ADN59771	Peptide-v
414	49	87.5	13	3	AAU17015	TPO-mimet	487	49	87.5	42	7	ADN59818	Peptide-v
415	49	87.5	13	5	ABB72901	TPO mimet	488	49	87.5	43	7	ADN59752	Peptide-v
416	49	87.5	13	7	ADJ73054	TPO mimet	489	49	87.5	46	7	ADN59789	Peptide-v
417	49	87.5	13	7	ADJ73052	TPO mimet	490	49	87.5	46	7	ADN59757	Peptide-v
418	49	87.5	13	7	ADJ73056	TPO mimet	491	49	87.5	46	7	ADN59783	Peptide-v
419	49	87.5	13	7	ADJ73053	TPO mimet	492	48	85.7	12	3	AAU17309	Synthetic
420	49	87.5	13	7	ADJ73055	TPO mimet	493	48	85.7	18	2	AAW09499	Thrombopo
421	49	87.5	13	8	ADJ52687	CH1 delet	494	48	85.7	18	2	AAW09459	Thrombopo
422	49	87.5	13	8	ADJ51648	CH1 delet	495	48	85.7	18	2	AAW36650	Thrombopo
423	49	87.5	14	3	AAU17017	TPO-mimet	496	48	85.7	18	2	AAW33026	Thrombopo
424	49	87.5	14	3	AAU17016	TPO-mimet	497	48	85.7	18	3	AAU17024	TPO-mimet
425	49	87.5	14	5	ABB72903	TPO mimet	498	48	85.7	18	4	AAU25869	Human thr
426	49	87.5	14	5	ABB72902	TPO mimet	499	48	85.7	18	4	AAU25823	Human thr
427	49	87.5	14	8	ADJ52689	CH1 delet	500	48	85.7	18	5	ABB72910	TPO mimet
428	49	87.5	14	8	ADJ52688	CH1 delet	501	48	85.7	18	7	ADJ73062	TPO mimet
429	49	87.5	14	8	ADJ51650	CH1 delet	502	48	85.7	18	7	ADN59675	Thrombopo
430	49	87.5	14	8	ADJ51649	CH1 delet	503	48	85.7	18	7	ADN59680	Thrombopo
431	49	87.5	15	3	AAU17018	TPO-mimet	504	48	85.7	18	7	ADN59661	Thrombopo
432	49	87.5	15	5	ABB72904	TPO mimet	505	48	85.7	18	7	ADN59657	Thrombopo
433	49	87.5	15	8	ADJ52691	CH1 delet	506	48	85.7	18	8	ADJ52697	CH1 delet
434	49	87.5	15	8	ADJ52690	CH1 delet	507	48	85.7	18	8	ADJ51658	CH1 delet
435	49	87.5	15	8	ADJ51652	CH1 delet	508	48	85.7	22	7	ADN59824	TMP pepti
436	49	87.5	15	8	ADJ51651	CH1 delet	509	48	85.7	22	7	ADN59828	TMP pepti
437	49	87.5	18	2	AAW09460	Thrombopo	510	48	85.7	23	7	ADN59775	Peptide-v
438	49	87.5	18	2	AAW09498	Thrombopo	511	48	85.7	23	7	ADN59793	Peptide-v
439	49	87.5	18	2	AAW36649	Thrombopo	512	48	85.7	25	7	ADN59696	Thrombopo
440	49	87.5	18	2	AAW33027	Thrombopo	513	48	85.7	25	7	ADN59704	Thrombopo
441	49	87.5	18	2	AAW36652	Thrombopo	514	48	85.7	25	7	ADN59730	Thrombopo
442	49	87.5	18	3	AAU17026	TPO-mimet	515	48	85.7	25	7	ADN59740	Thrombopo
443	49	87.5	18	4	AAU25868	Human thr	516	48	85.7	36	7	ADN59763	Peptide-v
444	49	87.5	18	4	AAU25824	Human thr	517	48	85.7	40	7	ADN59753	Peptide-v
445	49	87.5	18	4	AAU25871	Human thr	518	48	85.7	41	7	ADN59769	Peptide-v
446	49	87.5	18	5	ABB72912	TPO mimet	519	48	85.7	44	7	ADN59817	Peptide-v
447	49	87.5	18	7	ADJ73064	TPO mimet	520	48	85.7	46	7	ADN59781	Peptide-v
448	49	87.5	18	7	ADN59681	Thrombopo	521	48	85.7	46	7	ADN59787	Peptide-v
449	49	87.5	18	7	ADN59672	Thrombopo	522	47	83.9	8	2	AAW33037	Thrombopo
450	49	87.5	18	7	ADN59654	Thrombopo	523	47	83.9	8	4	AAU25982	Human thr
451	49	87.5	18	7	ADN59671	Thrombopo	524	47	83.9	8	5	ABP51678	TPO mimet
452	49	87.5	18	7	ADN59659	Thrombopo	525	47	83.9	8	5	ABP51678	TPO mimet
453	49	87.5	18	7	ADN59669	Thrombopo	526	47	83.9	8	9	ADQ16692	TPO mimet
454	49	87.5	18	7	ADN59667	Thrombopo	527	47	83.9	8	9	ADQ16692	TPO mimet
455	49	87.5	18	8	ADJ52699	CH1 delet	528	47	83.9	8	9	ADQ16693	TPO mimet
456	49	87.5	18	8	ADJ51660	CH1 delet	529	47	83.9	8	9	ADQ16692	TPO mimet
457	49	87.5	19	2	AAW09457	Thrombopo	530	47	83.9	18	9	ADV44431	Anti-teta
458	49	87.5	19	2	AAW09492	Thrombopo	531	47	83.9	18	9	ADV44431	Anti-teta
459	49	87.5	19	2	AAW36651	Thrombopo	532	46	82.1	14	8	ADN59730	Thrombopo
460	49	87.5	19	2	AAW33024	Thrombopo	533	46	82.1	14	8	ADN59730	Thrombopo
461	49	87.5	19	2	AAW36643	Thrombopo	534	46	82.1	15	4	AAW66723	Peptide C
462	49	87.5	19	3	AAU17021	TPO-mimet	535	46	82.1	15	4	AAU26028	Human thr
	49	87.5	19	4	AAU25862	Human thr		46	82.1	15	8	ADM72508	TPO mimet
	49	87.5	19	4	AAU25862	Human thr		46	82.1	18	7	ADN59658	Thrombopo

536	46	82.1	18	7	ADN59665	Adn59665 Thrombopo	609	45	80.4	14	4	AAU25866	Aau25866 Human thr
537	46	82.1	18	7	ADN59655	Adn59655 Thrombopo	610	45	80.4	14	4	AAU25988	Aau25988 Human thr
538	46	82.1	18	7	ADN59656	Adn59656 Thrombopo	611	45	80.4	14	5	ABB72900	Abb72900 TPO mimet
539	46	82.1	22	7	ADN59823	Adn59823 TMP pepti	612	45	80.4	14	7	ADJ73051	Adj73051 TPO mimet
540	46	82.1	22	7	ADN59822	Adn59822 TMP pepti	613	45	80.4	14	8	ADJ52686	Adj52686 CH1 delet
541	46	82.1	22	7	ADN59825	Adn59825 TMP pepti	614	45	80.4	14	8	ADJ51647	Adj51647 CH1 delet
542	46	82.1	22	7	ADN59832	Adn59832 TMP pepti	615	45	80.4	18	2	AAW09456	Aaw09456 Thrombopo
543	46	82.1	23	7	ADN59776	Adn59776 Peptide-v	616	45	80.4	18	2	AAW09487	Aaw09487 Thrombopo
544	46	82.1	23	7	ADN59774	Adn59774 Peptide-v	617	45	80.4	18	2	AAW33023	Aaw33023 Thrombopo
545	46	82.1	23	7	ADN59794	Adn59794 Peptide-v	618	45	80.4	18	2	AAW36638	Aaw36638 Thrombopo
546	46	82.1	23	7	ADN59694	Adn59694 Thrombopo	619	45	80.4	18	3	AAW17020	Aaw17020 TPO-mimet
547	46	82.1	25	7	ADN59694	Adn59694 Thrombopo	620	45	80.4	18	3	AAW36638	Aaw36638 Thrombopo
548	46	82.1	25	7	ADN59692	Adn59692 Thrombopo	621	45	80.4	18	4	AAU25857	Aau25857 Human thr
549	46	82.1	25	7	ADN59698	Adn59698 Thrombopo	622	45	80.4	18	4	AAU25820	Aau25820 Human thr
550	46	82.1	25	7	ADN59712	Adn59712 Thrombopo	623	45	80.4	18	5	ABB72906	Abb72906 TPO mimet
551	46	82.1	25	7	ADN59755	Adn59755 Peptide-v	624	45	80.4	18	7	ADJ73058	Adj73058 TPO mimet
552	46	82.1	25	7	ADN59762	Adn59762 Peptide-v	625	45	80.4	18	8	ADJ52693	Adj52693 CH1 delet
553	46	82.1	36	7	ADN59764	Adn59764 Peptide-v	626	45	80.4	18	8	ADJ51654	Adj51654 CH1 delet
554	46	82.1	36	7	ADN59764	Adn59764 Peptide-v	627	45	80.4	19	2	AAW09458	Aaw09458 Thrombopo
555	46	82.1	41	7	ADN59768	Adn59768 Peptide-v	628	45	80.4	19	2	AAW33025	Aaw33025 Thrombopo
556	46	82.1	41	7	ADN59770	Adn59770 Peptide-v	629	45	80.4	19	4	AAU25822	Aau25822 Human thr
557	46	82.1	41	7	ADN59768	Adn59768 Peptide-v	630	45	80.4	25	4	AAU26042	Aau26042 Human thr
558	46	82.1	46	7	ADN59786	Adn59786 Peptide-v	631	45	80.4	25	8	ADM72531	Adm72531 TPO mimet
559	46	82.1	46	7	ADN59782	Adn59782 Peptide-v	632	45	80.4	25	8	ADM72530	Adm72530 TPO mimet
560	46	82.1	46	7	ADN59788	Adn59788 Peptide-v	633	44	78.6	12	8	ADM72530	Adm72530 TPO mimet
561	45	80.4	11	2	AAW36791	Aaw36791 Thrombopo	634	44	78.6	13	4	AAU26041	Aau26041 Human thr
562	45	80.4	12	2	AAW35423	Aaw35423 Thrombopo	635	44	78.6	13	4	AAU26041	Aau26041 Human thr
563	45	80.4	12	2	AAU26000	Aau26000 Human thr	636	44	78.6	13	8	ADM72529	Adm72529 TPO mimet
564	45	80.4	12	2	AAW09467	Aaw09467 Thrombopo	637	44	78.6	14	2	AAW26040	Aaw26040 Human thr
565	45	80.4	13	2	AAW35399	Adn35399 Thrombopo	638	44	78.6	14	2	AAW09497	Aaw09497 Thrombopo
566	45	80.4	13	2	AAW35404	Adn35404 Thrombopo	639	44	78.6	18	2	AAW36653	Aaw36653 Thrombopo
567	45	80.4	13	2	AAW35417	Adn35417 Thrombopo	640	44	78.6	18	2	AAW36648	Aaw36648 Thrombopo
568	45	80.4	13	2	AAW33033	Adn33033 Thrombopo	641	44	78.6	18	3	AAW17025	Aaw17025 TPO-mimet
569	45	80.4	13	2	AAW35413	Adn35413 Thrombopo	642	44	78.6	18	3	AAU25872	Aau25872 Human thr
570	45	80.4	13	2	AAW35406	Adn35406 Thrombopo	643	44	78.6	18	4	AAU25867	Aau25867 Human thr
571	45	80.4	13	2	AAW35405	Adn35405 Thrombopo	644	44	78.6	18	5	ABB72911	Abb72911 TPO mimet
572	45	80.4	13	2	AAW35422	Adn35422 Thrombopo	645	44	78.6	18	7	ADJ73063	Adj73063 TPO mimet
573	45	80.4	13	2	AAW35397	Adn35397 Thrombopo	646	44	78.6	18	7	ADN59674	Adn59674 Thrombopo
574	45	80.4	13	2	AAW36783	Adn36783 Thrombopo	647	44	78.6	18	7	ADN59673	Adn59673 Thrombopo
575	45	80.4	13	4	AAU25994	Aau25994 Human thr	648	44	78.6	18	7	ADN59673	Adn59673 Thrombopo
576	45	80.4	13	4	AAU25997	Aau25997 Human thr	649	44	78.6	18	8	ADJ52698	Adj52698 CH1 delet
577	45	80.4	13	4	AAU25991	Aau25991 Human thr	650	44	78.6	18	8	ADJ51659	Adj51659 CH1 delet
578	45	80.4	13	4	AAU25984	Aau25984 Human thr	651	44	78.6	18	8	ADJ52698	Adj52698 CH1 delet
579	45	80.4	14	2	AAW09466	Aaw09466 Thrombopo	652	44	78.6	22	7	ADN59840	Adn59840 TMP pepti
580	45	80.4	14	2	AAW09462	Aaw09462 Thrombopo	653	44	78.6	22	7	ADN59841	Adn59841 TMP pepti
581	45	80.4	14	2	AAW09465	Aaw09465 Thrombopo	654	44	78.6	23	7	ADN59797	Adn59797 Peptide-v
582	45	80.4	14	2	AAW09482	Aaw09482 Thrombopo	655	44	78.6	23	7	ADN59778	Adn59778 Peptide-v
583	45	80.4	14	2	AAW33031	Adn33031 Thrombopo	656	44	78.6	23	7	ADN59779	Adn59779 Peptide-v
584	45	80.4	14	2	AAW35412	Adn35412 Thrombopo	657	44	78.6	23	7	ADN59796	Adn59796 Peptide-v
585	45	80.4	14	2	AAW35407	Adn35407 Thrombopo	658	44	78.6	25	7	ADN59726	Adn59726 Thrombopo
586	45	80.4	14	2	AAW35408	Adn35408 Thrombopo	659	44	78.6	35	7	ADN59754	Adn59754 Peptide-v
587	45	80.4	14	2	AAW35408	Adn35408 Thrombopo	660	44	78.6	35	7	ADN59754	Adn59754 Peptide-v
588	45	80.4	14	2	AAW36633	Adn36633 Thrombopo	661	44	78.6	36	7	ADN59766	Adn59766 Peptide-v
589	45	80.4	14	2	AAW35398	Adn35398 Thrombopo	662	44	78.6	36	7	ADN59756	Adn59756 Peptide-v
590	45	80.4	14	2	AAW33029	Adn33029 Thrombopo	663	44	78.6	36	7	ADN59767	Adn59767 Peptide-v
591	45	80.4	14	2	AAW35396	Adn35396 Thrombopo	664	44	78.6	36	7	ADN59767	Adn59767 Peptide-v
592	45	80.4	14	2	AAW35401	Adn35401 Thrombopo	665	44	78.6	41	7	ADN59816	Adn59816 Peptide-v
593	45	80.4	14	2	AAW35403	Adn35403 Thrombopo	666	44	78.6	41	7	ADN59772	Adn59772 Peptide-v
594	45	80.4	14	2	AAW36647	Adn36647 Thrombopo	667	44	78.6	41	7	ADN59773	Adn59773 Peptide-v
595	45	80.4	14	2	AAW35400	Adn35400 Thrombopo	668	44	78.6	46	7	ADN59790	Adn59790 Peptide-v
596	45	80.4	14	2	AAW35402	Adn35402 Thrombopo	669	44	78.6	46	7	ADN59785	Adn59785 Peptide-v
597	45	80.4	14	2	AAW33032	Adn33032 Thrombopo	670	43	76.8	46	7	ADN59784	Adn59784 Peptide-v
598	45	80.4	14	3	AAW17014	Aaw17014 TPO-mimet	671	43	76.8	46	7	ADN59791	Adn59791 Peptide-v
599	45	80.4	14	4	AAU25987	Aau25987 Human thr	672	43	76.8	11	2	AAW35425	Aaw35425 Thrombopo
600	45	80.4	14	4	AAU25826	Aau25826 Human thr	673	43	76.8	11	2	AAW35425	Aaw35425 Thrombopo
601	45	80.4	14	4	AAU25993	Aau25993 Human thr	674	43	76.8	14	8	ADM72507	Adm72507 TPO mimet
602	45	80.4	14	4	AAU25852	Aau25852 Human thr	675	43	76.8	14	8	ADM72519	Adm72519 TPO mimet
603	45	80.4	14	4	AAU25989	Aau25989 Human thr	676	43	76.8	14	8	ADM72505	Adm72505 TPO mimet
604	45	80.4	14	4	AAU25983	Aau25983 Human thr	677	43	76.8	14	8	ADM72475	Adm72475 TPO mimet
605	45	80.4	14	4	AAU25985	Aau25985 Human thr	678	43	76.8	14	8	ADM72501	Adm72501 TPO mimet
606	45	80.4	14	4	AAU25995	Aau25995 Human thr	679	43	76.8	14	9	ADU70205	Adu70205 Thrombopo
607	45	80.4	14	4	AAU25992	Aau25992 Human thr	680	43	76.8	14	9	ADU75978	Adu75978 Thrombopo
608	45	80.4	14	4	AAU25986	Aau25986 Human thr	681	43	76.8	14	9	ADY64334	Aay64334 Thrombopo
							681	43	76.8	15	2	AAW66722	Aaw66722 Peptide c

682	43	76.8	15	2	AAW66720	Peptide c	755	40	71.4	19	5	ABB72909	TPO mimet
683	43	76.8	15	2	AAW66728	Peptide c	756	40	71.4	19	7	ADJ73061	TPO mimet
684	43	76.8	15	2	AAW66728	Peptide c	757	40	71.4	19	8	ADJ52696	CHI delet
685	43	76.8	15	4	AAU26033	Human thr	758	40	71.4	19	8	ADJ51657	CHI delet
686	43	76.8	15	4	AAU25833	Human thr	759	40	71.4	329	4	ABG23127	Novel hum
687	43	76.8	15	4	AAU26027	Human thr	760	40	71.4	492	7	ABO82111	Pseudomon
688	43	76.8	15	4	AAU26025	Human thr	761	39	69.6	14	3	AAU26025	TPO mimet
689	43	76.8	15	4	AAU26029	Human thr	762	39	69.6	14	3	ABU72896	TPO mimet
690	43	76.8	15	8	ADM72510	TPO mimet	763	39	69.6	14	7	ADJ73047	TPO mimet
691	43	76.8	15	8	ADM72506	TPO mimet	764	39	69.6	14	8	ADJ52682	CHI delet
692	43	76.8	15	8	ADM72500	TPO mimet	765	39	69.6	14	8	ADJ51643	CHI delet
693	43	76.8	15	8	ADM72518	TPO mimet	766	39	69.6	95	5	AAU09166	Human cyt
694	43	76.8	15	8	ADM72477	TPO mimet	767	39	69.6	100	5	ABP47800	Human cyt
695	43	76.8	15	8	ADM72504	TPO mimet	768	39	69.6	168	4	ABB65448	Drosophil
696	43	76.8	15	9	ADY64335	Thrombopo	769	39	69.6	230	2	AAR60767	HIV-2 ROD
697	43	76.8	16	2	AAW66711	Peptide c	770	39	69.6	316	9	ABM93583	M. xanthu
698	43	76.8	18	8	ADM72476	TPO mimet	771	39	69.6	318	8	ADY22882	Plant ful
699	43	76.8	18	2	AAW09488	Thrombopo	772	39	69.6	428	5	ABP47757	Protein #
700	43	76.8	18	2	AAW36639	Thrombopo	773	39	69.6	450	3	AAG50584	Arabidops
701	43	76.8	18	4	AAU25858	Human thr	774	39	69.6	475	3	AAG50583	Arabidops
702	43	76.8	18	7	ADN59677	Thrombopo	775	39	69.6	499	3	AAG50582	Arabidops
703	43	76.8	25	7	ADN59734	Thrombopo	776	39	69.6	876	9	ABM92266	M. xanthu
704	43	76.8	28	8	ADJ52647	CHI delet	777	39	69.6	1014	1	AAP80810	Sequence
705	43	76.8	29	3	AAB16972	TPO mimet	778	39	69.6	1027	1	AAP81773	Sequence
706	43	76.8	29	5	ABB72858	TPO mimet	779	39	69.6	1035	2	AAR04025	Pol gene
707	43	76.8	29	7	ADJ73012	TPO mimet	780	39	69.6	1036	2	AAR20599	ROD HIV-2
708	43	76.8	29	8	ADJ73008	TPO mimet	781	39	69.6	1036	3	AAU51978	HIV-2 ROD
709	43	76.8	29	8	ADJ52643	CHI delet	782	39	69.6	1036	6	ADA07961	Pol
710	43	76.8	29	8	ADJ51604	CHI delet	783	39	69.6	1036	8	ADL23699	HIV-2 ROD
711	43	76.8	29	8	ADJ51608	CHI delet	784	38	67.9	10	2	AAW09473	Thrombopo
712	43	76.8	30	9	ADY64336	Thrombopo	785	38	67.9	10	2	AAW36624	Thrombopo
713	43	76.8	36	3	AAU96521	Cyclic or	786	38	67.9	10	4	AAU25843	Human thr
714	43	76.8	36	3	AAB17298	TPO-mimet	787	38	67.9	28	7	AAR38642	Human col
715	43	76.8	36	3	AAB17299	TPO-mimet	788	38	67.9	311	4	AAU41798	Human pol
716	43	75.0	18	2	AAW09592	Thrombopo	789	38	67.9	496	9	ABM91663	M. xanthu
717	42	75.0	18	2	AAW36743	Thrombopo	790	38	67.9	549	2	AAW40113	Human alp
718	42	75.0	18	4	AAU25962	Human thr	791	38	67.9	572	7	ABO80566	Pseudomon
719	42	75.0	36	3	AAU96522	Linear th	792	38	67.9	694	8	ADQ87654	Novel hum
720	42	75.0	36	3	AAU96522	Linear th	793	38	67.9	994	8	ABM82875	Human dia
721	42	75.0	100	5	ABP47801	Sooty man	794	38	67.9	1061	8	ABM82874	Human dia
722	42	75.0	429	5	ABP47758	Protein #	795	38	67.9	1078	4	AAU40012	Human pol
723	42	75.0	519	7	ADM05082	Human pro	796	38	67.9	1712	5	AAO17361	Human alp
724	41	73.2	18	7	ADN59679	Thrombopo	797	38	67.9	1712	8	ADL10473	Human the
725	41	73.2	18	7	ADN59676	Thrombopo	798	38	67.9	1712	8	ADU06526	Novel bro
726	41	73.2	18	7	ADN59662	Thrombopo	799	38	67.9	1712	9	ADY59885	Human col
727	41	73.2	22	7	ADN59829	TMP pepti	800	37.5	67.0	131	7	ABO67862	Pseudomon
728	41	73.2	25	7	ADN59738	Thrombopo	801	37.5	67.0	169	7	ABO76480	Pseudomon
729	41	73.2	25	7	ADN59706	Thrombopo	802	37	66.1	10	2	AAW35427	Thrombopo
730	41	73.2	25	7	ADN59732	Thrombopo	803	37	66.1	10	4	AAU26002	Human thr
731	41	73.2	40	7	ADN59760	Peptide-v	804	37	66.1	14	2	AAW09479	Thrombopo
732	41	73.2	302	6	ABU33828	Protein e	805	37	66.1	14	2	AAW09481	Thrombopo
733	41	73.2	308	4	AAU25854	Human thr	806	37	66.1	14	2	AAW36630	Thrombopo
734	40	71.4	14	2	AAW09484	C glutami	807	37	66.1	14	2	AAW36632	Thrombopo
735	40	71.4	14	2	AAW36766	Thrombopo	808	37	66.1	14	2	AAW36631	Thrombopo
736	40	71.4	14	2	AAW36635	Thrombopo	809	37	66.1	14	3	AAU17013	TPO-mimet
737	40	71.4	14	3	AAU17009	Thrombopo	810	37	66.1	14	3	AAU17013	TPO-mimet
738	40	71.4	14	4	AAU25854	Human thr	811	37	66.1	14	4	AAU25849	Human thr
739	40	71.4	14	5	ABB72895	TPO mimet	812	37	66.1	14	4	AAU25851	Human thr
740	40	71.4	14	7	ADJ73046	TPO mimet	813	37	66.1	14	5	ABB72899	TPO mimet
741	40	71.4	14	8	ADJ52681	CHI delet	814	37	66.1	14	7	ADJ73050	TPO mimet
742	40	71.4	14	8	ADJ52682	CHI delet	815	37	66.1	14	8	ADJ52685	CHI delet
743	40	71.4	14	8	ADM72513	TPO mimet	816	37	66.1	14	8	ADJ51646	CHI delet
744	40	71.4	15	2	AAW66725	Peptide c	817	37	66.1	14	8	ADM72515	TPO mimet
745	40	71.4	15	4	AAU25834	Human thr	818	37	66.1	14	8	ADM72521	TPO mimet
746	40	71.4	15	4	AAU26030	Human thr	819	37	66.1	15	2	AAW66726	Peptide c
747	40	71.4	15	8	ADM72512	TPO mimet	820	37	66.1	15	2	AAW66729	Peptide c
748	40	71.4	18	2	AAW09490	Thrombopo	821	37	66.1	15	4	AAU26031	Human thr
749	40	71.4	18	2	AAW36641	Thrombopo	822	37	66.1	15	4	AAU26034	Human thr
750	40	71.4	18	4	AAU25860	Human thr	823	37	66.1	15	8	ADM72520	TPO mimet
751	40	71.4	19	2	AAW09495	Thrombopo	824	37	66.1	15	8	ADM72514	TPO mimet
752	40	71.4	19	2	AAW36646	Thrombopo	825	37	66.1	132	7	ADM04387	Human pro
753	40	71.4	19	3	AAU17023	TPO-mimet	826	37	66.1	137	8	ABO84792	Murine ca
754	40	71.4	19	4	AAU25865	Human thr	827	37	66.1	180	7	ABO80306	Pseudomon

828	37	66.1	261	7	ADH86920	Adh86920	Enterococ	901	36	64.3	513	9	AEBA49395	AbE49395	N. gonorr
829	37	66.1	306	5	ABG91631	Abg91631	Purine/py	902	36	64.3	513	9	ABE49397	AbE49397	N. gonorr
830	37	66.1	306	5	ABG91624	Abg91624	Purine/py	903	36	64.3	520	7	ABO64867	AbO64867	Klebsiell
831	37	66.1	306	6	ABU36556	Abu36556	Protein e	904	36	64.3	592	4	ABG06780	AbG06780	Novel hum
832	37	66.1	306	6	ABU34427	Abu34427	Protein e	905	36	64.3	603	8	ADG25043	AdG25043	Bacterial
833	37	66.1	306	6	ABU35738	Abu35738	Protein e	906	36	64.3	607	3	AY58580	AY58580	Sorangium
834	37	66.1	321	8	ADK65923	Adk65923	Plant ful	907	36	64.3	685	6	ADA54712	AdA54712	Human pro
835	37	66.1	355	6	ABR58696	AbR58696	Human can	908	36	64.3	806	9	ABM93223	AbM93223	M. xanthu
836	37	66.1	355	6	ADF13714	Adf13714	Tumor-Ass	909	36	64.3	852	2	AAW79065	AAW79065	Feline Im
837	37	66.1	355	8	ADS00100	AdS00100	Human ocu	910	36	64.3	854	2	AAW44957	AAW44957	Feline Im
838	37	66.1	355	8	ADS00098	AdS00098	Human ocu	911	36	64.3	854	2	AAW1250	AAW1250	FIV PPR e
839	37	66.1	355	9	ADW14773	AdW14773	Tumor-ass	912	36	64.3	855	2	AAW28033	AAW28033	Env polyp
840	37	66.1	364	6	ABU23456	Abu23456	Protein e	913	36	64.3	855	2	AAW51252	AAW51252	FIV UK2 e
841	37	66.1	372	4	AAG80999	Ag80999	HIV prote	914	36	64.3	855	2	AAW51251	AAW51251	FIV JAPAN
842	37	66.1	476	8	ADT58335	AdT58335	Plant pol	915	36	64.3	855	2	AAW51255	AAW51255	FIV SWISS
843	37	66.1	491	8	ADG76665	AdG76665	Phospholi	916	36	64.3	855	2	AAW58585	AAW58585	FIV UK8 e
844	37	66.1	526	8	ADY07792	AdY07792	Plant ful	917	36	64.3	855	2	AAW23711	AAW23711	Feline im
845	37	66.1	777	8	ADS27768	AdS27768	Bacterial	918	36	64.3	856	2	AAW51249	AAW51249	FIV PET-F
846	37	66.1	861	9	ADW38367	AdW38367	Ancestral	919	36	64.3	856	2	AAW51253	AAW51253	FIV DUTCH
847	37	66.1	861	9	ADW38369	AdW38369	Ancestral	920	36	64.3	856	2	AAW51248	AAW51248	FIV petal
848	37	66.1	861	9	ADW38365	AdW38365	Ancestral	921	36	64.3	856	2	AAW51254	AAW51254	FIV envel
849	37	66.1	999	9	ADX40028	AdX40028	HIV pol p	922	36	64.3	856	2	AAW07846	AAW07846	Feline im
850	37	66.1	1055	14	AAW09478	AAW09478	Thrombopo	923	36	64.3	856	2	AAW08436	AAW08436	Feline im
851	36	64.3	14	2	AAW36629	AAW36629	Thrombopo	924	36	64.3	856	2	AAW01823	AAW01823	FIV envel
852	36	64.3	14	2	AAW36629	AAW36629	Thrombopo	925	36	64.3	856	2	AAW53671	AAW53671	FIV PPR c
853	36	64.3	14	4	AAU25848	AAU25848	Human thr	926	36	64.3	856	2	AAW53671	AAW53671	Ancestral
854	36	64.3	15	4	AAU26016	AAU26016	Human thr	927	36	64.3	856	2	ADW38353	AdW38353	Ancestral
855	36	64.3	15	8	ADM72480	AdM72480	TPO mimet	928	36	64.3	860	9	ADW38357	AdW38357	Ancestral
856	36	64.3	15	8	ADM72481	AdM72481	TPO mimet	929	36	64.3	861	9	ADW38359	AdW38359	Ancestral
857	36	64.3	15	9	ADM72481	AdM72481	TPO mimet	930	36	64.3	861	9	ADW38373	AdW38373	Ancestral
858	36	64.3	15	9	ADV21762	AdV21762	SIV pol p	931	36	64.3	861	9	ADW38371	AdW38371	Ancestral
859	36	64.3	15	9	ADV21761	AdV21761	SIV pol p	932	36	64.3	861	9	ADW38363	AdW38363	Ancestral
860	36	64.3	16	4	AAU26017	AAU26017	Human thr	933	36	64.3	861	9	ADW38375	AdW38375	Ancestral
861	36	64.3	18	2	AAW09489	AAW09489	Thrombopo	934	36	64.3	861	9	ADW38361	AdW38361	Ancestral
862	36	64.3	18	4	AAW36640	AAW36640	Thrombopo	935	36	64.3	861	9	ADW38355	AdW38355	Ancestral
863	36	64.3	110	4	AAW56555	AAW56555	Human thr	936	36	64.3	869	4	ABBS57877	ABBS57877	Drosophil
864	36	64.3	137	4	AAW91823	AAW91823	Human imm	937	36	64.3	870	2	AAW51246	AAW51246	Composite
865	36	64.3	142	8	ADW74465	AdW74465	Plant ful	938	36	64.3	999	9	ADX39926	AdX39926	HIV Pol p
866	36	64.3	169	7	ABM86829	AbM86829	Rice abio	939	36	64.3	1002	9	ADX40004	AdX40004	HIV Pol p
867	36	64.3	176	5	AAE25747	AAE25747	Rice FT h	940	36	64.3	1055	1	AAW13055	AAW13055	HIV-2 pro
868	36	64.3	176	8	ADR04263	AdR04263	Rice FT h	941	36	64.3	1056	1	AAW1783	AAW1783	Sequence
869	36	64.3	200	8	ADR23470	AdR23470	Bacterial	942	36	64.3	1060	2	AAW89314	AAW89314	Sequence
870	36	64.3	254	8	ADW74695	AdW74695	Plant ful	943	36	64.3	1060	2	AAW89314	AAW89314	Sequence
871	36	64.3	268	6	ABM70732	AbM70732	Phototrab	944	36	64.3	1072	4	ABW69797	ABW69797	Sequence
872	36	64.3	271	8	ADS42097	AdS42097	Bacterial	945	36	64.3	1118	9	ADV23765	AdV23765	Siv full
873	36	64.3	307	3	AAW51074	AAW51074	Arabidops	946	36	64.3	1119	8	ADQ21165	ADQ21165	Human sof
874	36	64.3	307	3	AAW51074	AAW51074	Arabidops	947	36	64.3	1150	5	ABW97549	ABW97549	Novel hum
875	36	64.3	312	5	ABW55368	ABW55368	Lactococc	948	36	64.3	1257	8	ADJ25467	ADJ25467	Neurocan
876	36	64.3	316	3	AAW14586	AAW14586	Arabidops	949	36	64.3	1257	8	ADJ25467	ADJ25467	Rat neuro
877	36	64.3	316	3	AAW14586	AAW14586	Arabidops	950	36	64.3	1275	6	ABJ25467	ABJ25467	Aspergill
878	36	64.3	316	5	ABP61149	ABP61149	Human GPC	951	36	64.3	1282	7	ABW85346	ABW85346	Mouse pro
879	36	64.3	316	7	ADC79356	AdC79356	Human G p	952	36	64.3	1307	6	ABW56962	ABW56962	Human AVA
880	36	64.3	324	5	ABW24556	ABW24556	Human olf	953	36	64.3	1335	7	ADC87653	ADC87653	Human mam
881	36	64.3	324	5	ABP61135	ABP61135	Human GPC	954	36	64.3	1335	7	ADC87655	ADC87655	Mouse mam
882	36	64.3	324	5	ABP95673	ABP95673	Human GPC	955	36	64.3	1335	7	ADC87655	ADC87655	Mouse mam
883	36	64.3	324	5	AAU95749	AAU95749	Human olf	956	36	64.3	1335	8	ADH18948	ADH18948	Human rap
884	36	64.3	324	5	AAU95176	AAU95176	G-coupled	957	36	64.3	1335	8	ADL91592	ADL91592	Human cel
885	36	64.3	324	5	ADC86379	AdC86379	Human GPC	958	36	64.3	1335	9	ADU40386	ADU40386	Human hum
886	36	64.3	324	8	ADG83490	AdG83490	Human Olf	959	36	64.3	1338	6	ABW56961	ABW56961	Human AVA
887	36	64.3	324	7	ADC79328	AdC79328	Human G p	960	36	64.3	1358	7	ABW53347	ABW53347	Human pro
888	36	64.3	325	7	AGS1072	AGS1072	Arabidops	961	36	64.3	1358	6	ABJ26067	ABJ26067	Aspergill
889	36	64.3	329	3	AAW14585	AAW14585	Arabidops	962	36	64.3	1435	6	ABJ26067	ABJ26067	Aspergill
890	36	64.3	331	3	AAW14585	AAW14585	Arabidops	963	36	64.3	1435	6	ABJ26067	ABJ26067	Aspergill
891	36	64.3	350	2	AAW62621	AAW62621	Mus muscu	964	35	62.5	10	2	AAW09471	AAW09471	Thrombopo
892	36	64.3	371	8	ADM98610	AdM98610	Geranylge	965	35	62.5	10	2	AAW36622	AAW36622	Thrombopo
893	36	64.3	408	5	AAE16177	AAE16177	Human G-p	966	35	62.5	10	2	AAW25841	AAW25841	Thrombopo
894	36	64.3	486	6	ABP77210	ABP77210	N. gonorr	967	35	62.5	18	2	AAW09595	AAW09595	Thrombopo
895	36	64.3	512	4	ABG25367	ABG25367	Novel hum	968	35	62.5	18	2	AAW36746	AAW36746	Thrombopo
896	36	64.3	513	2	AAW38789	AAW38789	Neisseria	969	35	62.5	18	2	AAU25965	AAU25965	Human thr
897	36	64.3	513	2	AAW38786	AAW38786	Neisseria	970	35	62.5	18	7	ADN59813	ADN59813	Thrombopo
898	36	64.3	513	2	AAW38787	AAW38787	Neisseria	971	35	62.5	18	7	ADN59814	ADN59814	Thrombopo
899	36	64.3	513	9	ABE49392	AbE49392	N. mening	972	35	62.5	79	4	ABM64243	ABM64243	Propionib
900	36	64.3	513	9	ABE49394	AbE49394	N. mening	973	35	62.5	79	6	ABM60762	ABM60762	Propionib

974 AAU63256 Propionib
975 Abm59775 Propionib
976 Abg07816 Novel hum
977 Abc068391 Pseudomon
978 Aag57785 Arabidops
979 Aag57786 Arabidops
980 Aag55516 Arabidops
981 Aag55791 Arabidops
982 Aag17852 Arabidops
983 Abu23180 Protein e
984 Aag17851 Arabidops
985 Aag55515 Arabidops
986 Aar99488 PopR. 10/
987 Abg22498 Novel hum
988 Aag56497 M. xanthu
989 Abu21946 Protein e
990 Aau19231 Human G p
991 Aau19230 Human G p
992 Add24938 Escherich
993 Aag17850 Arabidops
994 Abm59510 M. xanthu
995 Aag30100 Arabidops
996 Aag57790 Arabidops
997 Aag57789 Arabidops
998 Aag96815 A. thalia
999 Adt87825 Plant hom
1000 Aag59392 Arabidops

ALIGNMENTS

RESULT 1
AAB17006
ID AAB17006 standard; peptide; 10 AA.
XX AC AAB17006;
XX DT 31-OCT-2000 (first entry)
XX DE TPO-mimetic peptide sequence SEQ ID NO:62.
XX KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cycostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX OS Synthetic.
XX PN WO200024782-A2.
XX PD 04-MAY-2000.
XX PF 25-OCT-1999; 99WO-US025044.
XX PR 23-OCT-1998; 98US-0105371P.
XX PR 22-OCT-1999; 99US-00428082.
XX PA (AMGE-) AMGEN INC.
XX PF Feige U, Liu C, Cheetham J, Boone TC;
XX DR WPI; 2000-350702/30.
XX PT Novel composition of matter comprising an Fc domain and pharmacologically
XX active peptides, useful for treating cancer and autoimmune diseases.
XX P8 Claim 19; Page 216; 608pp; English.
XX CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-P1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAG59443 to AAG63526 and AAB19355 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.0042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLROWLA 10
Db 1 EGPTLROWLA 10
|||||

RESULT 2
AAB72892
ID AAB72892 standard; peptide; 10 AA.
XX AC ABE72892;
XX DT 05-APR-2002 (first entry)
XX DE TPO mimetic peptide SEQ ID NO:62.
XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cycostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200183525-A2.
XX PD 08-NOV-2001.
XX PF 02-MAY-2001; 2001WO-US014310.
XX PR 03-MAY-2000; 2000US-00563286.
XX PA (AMGE-) AMGEN INC.
XX PF Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX DR WPI; 2002-130313/17.
XX PT Novel vehicle-peptide molecule or its multimers useful for treating
XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
XX diabetic retinopathy, obesity, sleep disorders and infertility.

PS Claim 39; Page 43; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritis, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention

XX Sequence 10 AA;

Query Match 100.0%; Score 56; DB 5; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

RESULT 3

ID ADJ73043

ADJ73043 standard; peptide; 10 AA.

AC ADJ73043;

06-MAY-2004 (first entry)

TPO mimetic peptide sequence SeqID 497.

XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.

XX Synthetic.

OS WO2003084477-A2.

PN 16-OCT-2003.

PD 24-MAR-2003; 2003WO-US009139.

PF 29-MAR-2002; 2002US-0368791P.

PR (CENZ) CENTOCOR INC.

PI Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;

XX WPI; 2003-804237/75.

XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 497; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 10 AA;

Query Match 100.0%; Score 56; DB 7; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

RESULT 4

ADJ52678

ID ADJ52678 standard; peptide; 10 AA.

AC ADJ52678;

06-MAY-2004 (first entry)

CH1 deleted mimetibody-related peptide SeqID497.

XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.

XX Unidentified.

OS Synthetic.

PN WO2004002417-A2.

PD 08-JAN-2004.

PF 27-JUN-2003; 2003WO-US020347.

PR 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;

PI Kutolowski KA;

DR WPI; 2004-082870/08.

XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.

XX Claim 2; SEQ ID NO 497; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0042;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
 |||||
 Db 1 EGPTLRQWLA 10

RESULT 5

ADJ51639
 ID ADJ51639 standard; peptide; 10 AA.

XX AC ADJ51639;

DT 06-MAY-2004 (first entry)

XX CHI deleted mimetibody-related peptide SeqID497.

XX CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunologic disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX Unidentified.

OS Synthetic.

XX WO2004002424-A2.

XX 08-JAN-2004.

XX 30-JUN-2003; 2003WO-US020495.

XX 28-JUN-2002; 2002US-0392431P.

PR 19-SEP-2002; 2002US-0412144P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
 FI Kutoloeki KA;

XX WPI; 2004-082872/08.

XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic,
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX Claim 14; SEQ ID NO 497; 123pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, antiallergic, ophthalmological, neuroleptic, nephrotropic or
 CC antiinflammatory, neuroleptic, ophthalmological, hepatotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
 CC modulator or cytokine-agonist. The methods and compositions of the
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunologic, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX Sequence 10 AA;

Query Match 100.0%; Score 56; DB 8; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

RESULT 6

AAW36787
 ID AAW36787 standard; peptide; 12 AA.

XX AC AAW36787;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX Synthetic.

XX Key Location/Qualifiers

FT Cross-links 12

FT /note= "terminal carboxy group linked to epsilon amino
 group of Lys14 in AAW36787"

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US0009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

PT Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX Example 9; Page 78; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transplants. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX Sequence 12 AA;
 SQ

Query Match 100.0%; Score 56; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGPTLRQWLA 10
 |||||
 DB 1 EGPTLRQWLA 10

RESULT 7
 AAW36781
 ID AAW36781 standard; peptide; 12 AA.
 XX
 AC AAW36781;
 DT 11-MAR-1998 (first entry)
 DE Thrombopoietin receptor binding peptide.
 DE
 XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopenia; chemotherapy;
 KW radiation therapy; bone marrow transduction; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Cross-links 12
 FT /note= "terminal carboxy group linked to epsilon amino
 FT group of Lys14 in AAW36782"
 FT
 XX WO9640750-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX
 XX 07-JUN-1996; 96WO-US009623.
 PF
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR
 XX 07-JUN-1995; 95US-00485301.
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 PI WPI; 1997-052226/05.
 XX
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
 XX Example 9; Page 78; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transplants. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX Sequence 12 AA;
 SQ

Query Match 100.0%; Score 56; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EGPTLRQWLA 10
 |||||
 DB 2 EGPTLRQWLA 11

RESULT 8
 AAU26014
 ID AAU26014 standard; peptide; 12 AA.
 XX
 AC AAU26014;
 XX
 DT 17-DEC-2001 (first entry)
 DE Human thrombopoietin receptor (TPO-R) activator peptide #200.
 DE
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX Homo sapiens.
 OS
 XX US6251864-B1.
 PN
 XX 26-JUN-2001.
 PD
 XX
 XX 01-MAR-2000; 2000US-00516704.
 PF
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR
 XX 07-JUN-1995; 95US-00485301.
 PR
 XX 07-JUN-1996; 96WO-US009623.
 PR
 XX 15-AUG-1996; 96US-00699027.
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 DR
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX Disclosure; Col 151; 128pp; English.
 PS
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living

CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 56; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

RESULT 9

AAW36792

ID AAW36792 standard; peptide; 13 AA.

XX AC AAW36792;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "optionally acylated"

FT Modified-site 12 /label= bala

FT Cross-links 13

FT /note= "epsilon amino group of Lys13 linked to terminal
FT carboxy group of AAW36791"

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 78; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of

CC thrombopoietin dependent cell lines

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 56; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 1 EGPTLRQWLA 10

RESULT 10

AAW36779

ID AAW36779 standard; peptide; 13 AA.

XX AC AAW36779;

XX 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Cross-links 13

FT /note= "terminal carboxy group linked to epsilon amino
FT group of Lys15 in AAW36780"

XX WO9640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

XX Example 9; Page 78; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 56; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

```

Db      ||||| 2 EGPTLRQWLA 11
        ||||| 10
        ||||| 11

RESULT 11
AAU26008
ID AAU26008 standard; peptide; 13 AA.
XX
AC AAU26008;
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #194.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
XX Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
XX Yin Q;
XX
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
XX with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 147; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX of activating thrombopoietin receptors in cells comprise contacting the
XX cells with effective amounts of peptides and peptide mimetics attached to
XX hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX as that due to chemotherapy, radiation therapy or bone-marrow
XX transplantation and to prevent thrombocytopenia in patients at risk. The
XX sequences are used to treat and prevent haematological disorders
XX including thrombocytopenia and platelet disorders. They are used in vitro
XX as unique tools for understanding the biological role of thrombopoietin
XX (TPO) and to develop other compounds that bind to and activate the TPO
XX receptor. The peptides can be used to detect TPO receptors on living
XX cells and fixed cells, in biological fluids, in tissue homogenates, and
XX in purified or natural biological materials. They may also be used for in
XX situ staining, fluorescence-activated cell sorting, Western blotting and
XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX be used for in vitro expansion of megakaryocytes and their committed
XX progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 56; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      ||||| 2 EGPTLRQWLA 11
        ||||| 10
        ||||| 11

RESULT 12
AAU26012
ID AAU26012 standard; peptide; 13 AA.
XX
AC AAU26012;
DT 17-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #198.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
XX Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
XX Yin Q;
XX
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
XX with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 149-150; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX of activating thrombopoietin receptors in cells comprise contacting the
XX cells with effective amounts of peptides and peptide mimetics attached to
XX hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX as that due to chemotherapy, radiation therapy or bone-marrow
XX transplantation and to prevent thrombocytopenia in patients at risk. The
XX sequences are used to treat and prevent haematological disorders
XX including thrombocytopenia and platelet disorders. They are used in vitro
XX as unique tools for understanding the biological role of thrombopoietin
XX (TPO) and to develop other compounds that bind to and activate the TPO
XX receptor. The peptides can be used to detect TPO receptors on living
XX cells and fixed cells, in biological fluids, in tissue homogenates, and
XX in purified or natural biological materials. They may also be used for in
XX situ staining, fluorescence-activated cell sorting, Western blotting and
XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX be used for in vitro expansion of megakaryocytes and their committed
XX progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 56; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 EGPTLQWLA 10
Db      1 EGPTLQWLA 10
|||||
|

RESULT 13
AAU26015
ID AAU26015 standard; peptide; 13 AA.
AC AAU26015;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human thrombopoietin receptor (TPO-R) activator peptide #201.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ,
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 151; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 13 AA;
Query Match 100.0%; Score 56; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EGPTLQWLA 10
Db      1 EGPTLQWLA 10
|||||
|

RESULT 14
ADU73003
ID ADU73003 standard; peptide; 13 AA.
XX
AC ADU73003;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 457.
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
XX WO2003084477-A2.
XX
XX 16-OCT-2003.
XX
XX 24-MAR-2003; 2003WO-US009139.
XX
XX 29-MAR-2002; 2002US-0368791P.
PR
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX WPI; 2003-804237/75.
XX
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
XX Disclosure; SEQ ID NO 457; 97pp; English.
XX
XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
XX Sequence 13 AA;
Query Match 100.0%; Score 56; DB 7; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 EGPTLQWLA 10
Db      1 EGPTLQWLA 10
|||||
|

```

RESULT 15
ADJ52638
ID ADJ52638 standard; peptide; 13 AA.
XX AC
XX ADJ52638;
XX AC
XX DT 06-MAY-2004 (first entry)
XX XX
XX CH1 deleted mimetibody-related peptide SeqID457.
XX DE
XX KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
XX KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
XX KW fungicide; gene therapy; immune disorder; cardiovascular disease;
XX KW arrhythmia; hypertension; heart failure; neurodegenerative;
XX KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
XX KW cancerous condition; infectious disease; bacterial infection;
XX KW viral infection; fungal infection.
XX OS Unidentified.
XX OS Synthetic.
XX WO2004002417-A2.
XX PN
XX PD 08-JAN-2004.
XX PF 27-JUN-2003; 2003WO-US020347.
XX PR 28-JUN-2002; 2002US-0392431P.
XX XX (CENZ) CENTOCOR INC.
XX PA
XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX PI Kutoloski KA;
XX DR WPI; 2004-082870/08.
XX XX
XX PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
XX PT modulating, treating, alleviating, preventing an immune, cardiovascular,
XX PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
XX PT diseases.
XX PS Claim 2; SEQ ID NO 457; 129pp; English.
XX XX
XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
XX CC which encode them), compositions, methods and uses. The invention may be
XX CC useful for the development of compounds with an immunosuppressive,
XX CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
XX CC antibacterial, virucide or fungicide activity. In addition, the disclosed
XX CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
XX CC is useful for diagnosing or treating a disease condition in a cell,
XX CC tissue, organ or animal, specifically for modulating, treating,
XX CC alleviating, preventing the incidence or reducing the symptoms of an
XX CC immune, cardiovascular (for example arrhythmia, hypertension or heart
XX CC failure), or neurodegenerative (for example multiple sclerosis, dementia
XX CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
XX CC conditions, or infectious diseases (for example bacterial, viral or
XX CC fungal infection). The present sequence is that of a peptide which may be
XX CC used during the creation of a mimetibody of the invention.
XX SQ Sequence 13 AA;

XX Query Match 100.0%; Score 56; DB 8; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 0.0055;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTRLQWL 10
DB 1 EGPTRLQWL 10

RESULT 16
ADJ51599
ID ADJ51599 standard; peptide; 13 AA.

XX AC ADJ51599;
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE CH1 deleted mimetibody-related peptide SeqID457.
XX KW
XX KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
XX KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
XX KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
XX KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
XX KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
XX KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
XX KW dental disorder; oral disorder; dermatological disorder; ear disorder;
XX KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
XX KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
XX KW obstetric disorder; haematologic disorder; immunologic disorder;
XX KW allergic disorder; infectious disorder; musculoskeletal disorder;
XX KW oncological disorder; neurological disorder; nutritional disorder;
XX KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
XX KW renal disorder; pulmonary disorder.
XX XX Unidentified.
XX OS Synthetic.
XX WO2004002424-A2.
XX PN
XX PD 08-JAN-2004.
XX PF 30-JUN-2003; 2003WO-US020495.
XX PR 28-JUN-2002; 2002US-0392431P.
XX PR 19-SEP-2002; 2002US-0412144P.
XX XX (CENZ) CENTOCOR INC.
XX PA
XX PI Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX PI Kutoloski KA;
XX DR WPI; 2004-082872/08.
XX XX
XX PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
XX PT diagnosing, preventing or treating cardiovascular, dermatologic,
XX PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
XX PT nutritional disorders.
XX PS Claim 14; SEQ ID NO 457; 123pp; English.
XX XX
XX CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
XX CC which encode them), compositions, methods and uses. The invention may be
XX CC useful for the development of compounds with an osteopathic,
XX CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
XX CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
XX CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
XX CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
XX CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
XX CC modulator or cytokine-agonist. The methods and compositions of the
XX CC present invention are useful for the diagnosis, prevention and/or
XX CC treatment of diseases or conditions associated with aberrant expression
XX CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
XX CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
XX CC endocrine, haematologic, gastrointestinal, allergic, infectious,
XX CC obstetric, metabolic, gynaecological, immunologic, ophthalmologic,
XX CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
XX CC pediatric, psychiatric, renal or pulmonary disorders. The present
XX CC sequence is that of a peptide which may be used during the creation of a
XX CC mimetibody of the invention.
XX SQ Sequence 13 AA;

XX Query Match 100.0%; Score 56; DB 8; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 0.0055;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
 |||||
 Db 1 EGPTLRQWLA 10

RESULT 17
 AAW09463
 ID AAW09463 standard; protein; 14 AA.
 AC AAW09463;
 XX
 XX 10-SEP-1997 (first entry)
 DE Thrombopoietin receptor binding compound peptide.
 XX
 XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 1. 14
 FT /note= "Preferably linkages are selected from: -
 FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
 FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
 FT lower alkyl"
 FT Modified-site 1
 FT /note= "Preferably N-terminus is selected from: -NRR1; -
 FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
 FT benzoyloxycarbonyl-NH; benzoyloxycarbonyl-NH with 1-3
 FT substitutions on the phenyl ring selected from lower
 FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
 FT independently selected from hydrogen and lower alkyl"
 FT Modified-site 14
 FT /note= "Preferably C-terminus is -C(O)R2 where R2 is
 FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
 FT and R4 are independently selected from hydrogen and lower
 FT alkyl, and where the nitrogen atom of the -NR3R4 group
 FT can optionally be the amine group of the N-terminus of
 FT the peptide forming a cyclic peptide"

MO9640189-A1.
 19-DEC-1996.
 05-JUN-1996; 96WO-US008998.
 07-JUN-1995; 95US-00472371.
 07-JUN-1995; 95US-00473604.
 07-JUN-1995; 95US-00476168.
 07-JUN-1995; 95US-00478128.
 07-JUN-1995; 95US-00484090.
 07-JUN-1995; 95US-00485301.
 (GLAX) GLAXO GROUP LTD.
 Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
 Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 WPI; 1997-051883/05.
 Thrombopoietin receptor-binding/activating peptide(s) and peptide
 mimetic(s) - useful in treatment of haematological disorders, esp.
 thrombocytopenia resulting from chemotherapy, etc.
 Claim 18; Page 89; 106pp; English.
 The present sequence is a compound which binds to thrombopoietin (TPO)
 receptor (TR). It has a molecular weight of < 8000 Da, and a binding
 affinity to TR as expressed by an IC50 of no more than about 100 nm. The
 compound (especially if modified, see features table) can be used for
 treating patients suffering from haematological disorders and

CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. The peptide may also be used to maintain the
 CC proliferation and growth of TPO-dependent cell lines and for use in
 CC biological research, for detecting TPO receptors on living cells
 XX
 SQ Sequence 14 AA;
 Query Match 100.0%; Score 56; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 EGPTLRQWLA 10
 |||||
 Db 2 EGPTLRQWLA 11

RESULT 18
 AAW09468
 ID AAW09468 standard; protein; 14 AA.
 AC AAW09468;
 XX
 XX 10-SEP-1997 (first entry)
 DE Thrombopoietin receptor binding compound peptide (part of a dimer).
 XX
 XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 KW bone marrow transfusion; chemotherapy; radiation therapy.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Cross-links 14
 FT /note= "Linked to the omega Lys from AAW19534"
 FT MO9640189-A1.
 19-DEC-1996.
 05-JUN-1996; 96WO-US008998.
 07-JUN-1995; 95US-00472371.
 07-JUN-1995; 95US-00473604.
 07-JUN-1995; 95US-00476168.
 07-JUN-1995; 95US-00478128.
 07-JUN-1995; 95US-00484090.
 07-JUN-1995; 95US-00485301.
 (GLAX) GLAXO GROUP LTD.
 Dower WJ, Barrett RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
 Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 WPI; 1997-051883/05.
 Thrombopoietin receptor-binding/activating peptide(s) and peptide
 mimetic(s) - useful in treatment of haematological disorders, esp.
 thrombocytopenia resulting from chemotherapy, etc.
 Claim 30; Page 91; 106pp; English.
 The present sequence is a compound which binds to thrombopoietin (TPO)
 receptor (TR). It is part of a dimer linked by the omega amino acid to
 the omega amino acid in the sequence in AAW19534. The compound can be
 used for treating patients suffering from haematological disorders and
 CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. The peptide may also be used to maintain the
 CC proliferation and growth of TPO-dependent cell lines and for use in
 CC biological research, for detecting TPO receptors on living cells
 XX
 SQ Sequence 14 AA;
 Query Match 100.0%; Score 56; DB 2; Length 14;
 Query Match

Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTRLQWLA 10
Db 2 EGPTRLQWLA 11

RESULT 19
AAW33030
ID AAW33030 standard; peptide; 14 AA.
AC AAW33030;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
PS Claim 19; Page 89; 106pp; English.
XX
CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopaenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTRLQWLA 10
Db 2 EGPTRLQWLA 11

RESULT 20
AAW33034
ID AAW33034 standard; peptide; 14 AA.
XX

AC AAW33034;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
DE Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cross-links 14
FT /note= "terminal carboxy group linked to epsilon amino
FT group of Lys16 in AAW33035"
XX
PN WO9640750-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009623.
XX
PR 07-JUN-1995; 95US-00478128.
XX
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
WPI; 1997-052226/05.
XX
PT Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
PS Claim 30; Page 91; 106pp; English.
XX
CC The present peptide binds the thrombopoietin receptor (TR), has a
CC molecular weight of less than 8000 Da and a TR binding affinity as
CC expressed by an IC50 of no more than about 100 microm. It can be used to
CC treat disorders which are susceptible to treatment with a thrombopoietin
CC agonist, preferably haematological disorders and thrombocytopaenia
CC resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. It can also be used diagnostically, e.g. to investigate the
CC mechanism of thrombopoietin signal transduction and receptor activation,
CC or to maintain the proliferation and growth of thrombopoietin dependent
CC cell lines
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 56; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTRLQWLA 10
Db 2 EGPTRLQWLA 11

RESULT 21
AAW36782
ID AAW36782 standard; peptide; 14 AA.
XX
AC AAW36782;
XX
DT 11-MAR-1998 (first entry)
XX
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US0009623.
XX
PR 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
XX
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
PT thrombopoietin receptor - useful in treatment of haematological
PT disorders, esp. thrombocytopenia resulting from chemotherapy, etc.
XX
XX Example 9; Page 78; 106pp; English.
XX
XX The present peptide, which binds the thrombopoietin receptor (TR), can be
CC used to treat disorders which are susceptible to treatment with a
CC thrombopoietin agonist, preferably haematological disorders and
CC thrombocytopenia resulting from chemotherapy, radiation therapy or bone
CC marrow transfusions. It can also be used diagnostically, e.g. to
CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 56; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EGPTLRQWLA 10
XX |||||
XX Db 1 EGPTLRQWLA 10
XX
XX RESULT 24
XX ADI24843
XX ID ADI24843 standard; peptide; 14 AA.
XX
XX AC ADI24843;
XX
XX DT 15-APR-2004 (first entry)
XX
XX DE AF 12505 as active moiety for pharmacologically active peptide.
XX
XX KW pharmacologically active peptide conjugate; enzymatic cleavage; pain;
XX HIV; cancer; diabetes; incontinence; hypertension; amnesia;
XX Alzheimer's disease; fever; depression; sex hormone regulation;
XX eating disorder; schizophrenia; osteoporosis; insomnia;
XX Central nervous system disorder; contraceptive.
XX
XX OS Synthetic.
XX
XX PN WO9946283-A1.
XX
XX PD 16-SEP-1999.
XX
XX PF 09-MAR-1999; 99WO-DK000118.
XX
XX PR 09-MAR-1998; 98DK-00000317.
XX
XX PA (ZEAL-) ZEALAND PHARM AS.
XX
XX PI Larsen BD;
XX
XX DR WPI; 1999-561659/47.
XX
XX PT New peptide conjugates used for treating, e.g. pain, HIV, depression,
XX schizophrenia, osteoporosis or insomnia.
XX

PS Claim 24; Page 90; 113pp; English.
XX
XX The invention relates to a novel pharmacologically active peptide
CC conjugate having a reduced tendency towards enzymatic cleavage comprises
CC X and Z, where: (a) X is a pharmacologically active peptide sequence; and
CC (b) Z is a stabilising peptide sequence of 4-20 amino acid units
CC covalently bound to X, where each amino acid unit in the stabilizing
CC peptide sequence, Z being selected from Ala, Leu, Ser, Thr, Tyr, Asn,
CC Gln, Asp, Glu, Lys, Arg, His, Met, Orn, and amino acid units of formula -
CC NH-C(R1)(R2)-C(=O)-(1), where: R1 and R2 are H, 1-6C alkyl, phenyl, and
CC phenyl-methyl, where 1-6C-alkyl is optionally substituted with 1-3
CC substituents selected from halogen, hydroxy, amino, cyano, nitro.
CC sulfonyl, and carboxy, and phenyl and phenyl-methyl are optionally
CC substituted with 1-3 substituents selected from 1-6C-alkyl, 2-6C-alkenyl,
CC halogen, hydroxy, amino, cyano, nitro, sulfonyl, and carboxy, or R1 and R
CC 2 together with the C atom to which they are bound form a cyclopentyl,
CC cyclohexyl or cycloheptyl ring, e.g. 2,4-diaminobutanoic acid and 2,3-
CC diaminopropanoic acid; the ratio between the half-life of the peptide
CC conjugate and the half-life of the corresponding active peptide sequence,
CC X, when treated with carboxypeptidase A or leucine aminopeptidase in
CC about 50 mM phosphate buffer solution at about pH 7.4 and about 37 deg C
CC or in serum or plasma is at least about 2 (preferably at least about 10),
CC or when the pharmacologically active peptide X is not orally absorbed,
CC the conjugate is adsorbed, or a salt, with the proviso that the
CC pharmacologically active peptide conjugate is not selected from sequences
CC (ADI24837)-(ADI24841). The peptide conjugates can be used for treating
CC e.g. pain, HIV, cancer, diabetes, incontinence, hypertension, amnesia,
CC Alzheimer's disease, fever, depression, sex hormone regulation, eating
CC disorders, schizophrenia, osteoporosis or insomnia. They can also be used
CC for treating e.g. CNS disorders and as contraceptives. The conjugated
CC peptides are less susceptible to degradation by proteases compared to the
CC corresponding free pharmacologically active peptides. This sequence
CC represents a pharmacologically active peptide as the X part of the
CC peptide of the invention.
XX
XX Sequence 14 AA;
XX
XX Query Match 100.0%; Score 56; DB 2; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 EGPTLRQWLA 10
XX |||||
XX Db 2 EGPTLRQWLA 11
XX
XX RESULT 25
XX AAY96515
XX ID AAY96515 standard; peptide; 14 AA.
XX
XX AC AAY96515;
XX
XX DT 04-SEP-2000 (first entry)
XX
XX DE Thrombopoietin mimetic peptide.
XX
XX KW Thrombopoietin; mimetic; TMP; platelet; megakaryocyte; production;
XX anti-human immunodeficiency virus; anti-HIV; anti-anemic; dermatological;
XX immunosuppressive; anti-inflammatory.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX Modified-site 14
XX /note= "subunits in the dimer are covalently bonded at
XX each carboxy terminus through peptide linkage with NH2-
XX (CH2)4-CH(CONH2)-NH-CO-(CH2)2-NH2"
XX
XX PN WO200024770-A2.
XX
XX PD 04-MAY-2000.
XX
XX PF 22-OCT-1999; 99WO-US024834.
XX

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically

XX active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 203; 608pp; English.

XX The present invention describes composition of matter (I) comprising an

XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

XX (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each

XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-

XX (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

XX P3, and P4 = are each independently sequences of pharmacologically active

XX peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

XX c, d, e, and f = are each independently 0 or 1, provided that at least 1

XX of a and b is 1. The composition can have cytostatic, antiasthmatic,

XX thrombolytic and immunosuppressive activities. DNAs, vectors and host

XX cells from the present invention can be used for producing pharmaceutical

XX compositions. The compositions are useful for treating cancer, asthma,

XX thrombosis, or autoimmune diseases. The use of an Fc domain (rather than

XX such as Fc receptor binding, protein A binding, complement fixation, and

XX possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to

XX AAB18003 represent nucleotide and amino acid sequences used in the

XX exemplification of the present invention

XX Sequence 14 AA;

Query Match 100.0%; Score 56; DB 3; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 2 EGPTLRQWLA 11

RESULT 28

AAU26009

ID AAU26009 standard; peptide; 14 AA.

AC AAU26009;

XX 18-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #195.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

XX bone marrow transplantation; haematological disorder; platelet disorder;

XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;

XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;

XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

XX Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

XX Yin Q;

XX WPI; 2001-564142/63.

XX Activating thrombopoietin receptors in cells, used to treat

XX thrombocytopenia and hematological disorders, comprises contacting cells

XX with peptides and peptide mimetics attached to hydrophilic polymers.

XX Disclosure; Col 149-150; 128pp; English.

XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that

XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods

XX of activating thrombopoietin receptors in cells comprise contacting the

XX cells with effective amounts of peptides and peptide mimetics attached to

XX hydrophilic polymers. The methods are used to treat thrombocytopenia such

XX as that due to chemotherapy, radiation therapy or bone-marrow

XX transplantation and to prevent thrombocytopenia in patients at risk. The

XX sequences are used to treat and prevent haematological disorders

XX including thrombocytopenia and platelet disorders. They are used in vitro

XX as unique tools for understanding the biological role of thrombopoietin

XX (TPO) and to develop other compounds that bind to and activate the TPO

XX receptor. The peptides can be used to detect TPO receptors on living

XX cells and fixed cells, in biological fluids, in tissue homogenates, and

XX in purified or natural biological materials. They may also be used for in

XX situ staining, fluorescence-activated cell sorting, Western blotting and

XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can

XX be used for in vitro expansion of megakaryocytes and their committed

XX progenitors alone or in conjunction with additional cytokines

XX Sequence 14 AA;

Query Match 100.0%; Score 56; DB 4; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 2 EGPTLRQWLA 11

RESULT 29

AAU26013

ID AAU26013 standard; peptide; 14 AA.

AC AAU26013;

XX 18-DEC-2001 (first entry)

XX Human thrombopoietin receptor (TPO-R) activator peptide #199.

XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;

XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;

XX bone marrow transplantation; haematological disorder; platelet disorder;

XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;

XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;

XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX Homo sapiens.

XX US6251864-B1.

XX 26-JUN-2001.

XX 01-MAR-2000; 2000US-00516704.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX 07-JUN-1996; 96WO-US009623.

XX 15-AUG-1996; 96US-00699027.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PJ;

XX Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;

```

PI Yin Q;
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 151-152; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent hematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGPTLRQWLA 10
DB 1 EGPTLRQWLA 10
RESULT 30
AAU26006
ID AAU26006 standard; peptide; 14 AA.
AC AAU26006;
XX
XX 17-DEC-2001 (first entry)
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #192.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.
XX
XX Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirle SE, Gates CM, Schatz PJ;
PI

```

```

PI Yin Q;
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 151-152; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
CC progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 56; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EGPTLRQWLA 10
DB 1 EGPTLRQWLA 10
RESULT 30
AAU25827
ID AAU25827 standard; peptide; 14 AA.
AC AAU25827;
XX
XX 17-DEC-2001 (first entry)
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #13.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.
XX
XX Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
XX
XX 07-JUN-1995; 95US-00485301.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 15-AUG-1996; 96US-00699027.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwirle SE, Gates CM, Schatz PJ;
PI

```

PI Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 DR
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 XX Disclosure; Col 69-70; 128pp; English.
 PS
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 XX Sequence 14 AA;
 SQ

Query Match 100.0%; Score 56; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTLRQWLA 10
 DB 2 EGPTLRQWLA 11
 RESULT 32
 AAU26010
 ID AAU26010 standard; peptide; 14 AA.
 AC
 XX AAU26010;
 DT 17-DEC-2001 (first entry)
 XX Human thrombopoietin receptor (TPO-R) activator peptide #196.
 DE
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX
 XX Homo sapiens.
 OS
 XX US6251864-B1.
 PN
 XX 26-JUN-2001.
 PD
 XX 01-MAR-2000; 2000US-00516704.
 PF
 XX 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US009623.
 PR 15-AUG-1996; 96US-00699027.
 XX
 XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;
 XX WPI; 2001-564142/63.
 DR
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 XX Disclosure; Col 149-150; 128pp; English.
 PS
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 100.0%; Score 56; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTLRQWLA 10
 DB 1 EGPTLRQWLA 10
 RESULT 33
 AAU26004
 ID AAU26004 standard; peptide; 14 AA.
 AC
 XX AAU26004;
 DT 17-DEC-2001 (first entry)
 XX Human thrombopoietin receptor (TPO-R) activator peptide #190.
 DE
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX
 XX Homo sapiens.
 OS
 XX US6251864-B1.
 PN
 XX 26-JUN-2001.
 PD
 XX 01-MAR-2000; 2000US-00516704.
 PF
 XX 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 PR 07-JUN-1996; 96WO-US009623.
 PR 15-AUG-1996; 96US-00699027.
 XX
 XX (GLAX) GLAXO GROUP LTD.

PA (GLAXO) GLAXO GROUP LTD.
 XX Dower WJ, Barrett RW, Cwirila SE, Gates CM, Schatz PU;
 PI Balasubramanian P, Wegstrom CR, Hendren RW, Deprince RB, Podduturi S;
 XX Yin Q;
 DR WPI; 2001-564142/63.
 XX
 XX Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.
 XX
 XX Disclosure; Col 147; 128pp; English.
 PS
 XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 100.0%; Score 56; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTLRQWLA 10
 Db |||||||
 2 EGPTLRQWLA 11
 RESULT 34
 ABB72854
 ID ABB72854 standard; peptide; 14 AA.
 XX
 AC ABB72854;
 XX
 XX 05-APR-2002 (first entry)
 XX
 DE TPO mimetic peptide SEQ ID NO:24.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytosatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200183525-A2.
 XX
 XX 08-NOV-2001.
 PD

XX 02-MAY-2001; 2001WO-US014310.
 XX
 PR 03-MAY-2000; 2000US-00563286.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 XX Feige U, Liu C, Cheestham JC, Boone TC, Gudas JM;
 PI WPI; 2002-130313/17.
 DR
 XX Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 XX
 XX Claim 39; Page 43; 176pp; English.
 PS
 XX The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytosatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 100.0%; Score 56; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTLRQWLA 10
 Db |||||||
 2 EGPTLRQWLA 11
 RESULT 35
 ABB72853
 ID ABB72853 standard; peptide; 14 AA.
 XX
 AC ABB72853;
 XX
 XX 05-APR-2002 (first entry)
 XX
 DE TPO mimetic peptide SEQ ID NO:13.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytosatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200183525-A2.
 XX
 XX 08-NOV-2001.
 PD

XX OS Homo sapiens.
 OS Synthetic.
 XX PN WO200183525-A2.
 XX PN 08-NOV-2001.
 XX PD
 XX PF
 XX PR 02-MAY-2001; 2001WO-US014310.
 XX PR 03-MAY-2000; 2000US-00563286.
 XX PR (AMGE-) AMGEN INC.
 XX PA
 XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX PI WPI; 2002-130313/17.
 XX DR
 XX PT Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.
 XX PT
 XX PS Claim 39; Page 43; 176pp; English.
 XX CC The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antifertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention
 XX CC
 XX SQ Sequence 14 AA;
 Query Match 100.0%; Score 56; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTRLQWL A 10
 DB |||||
 2 EGPTRLQWL A 11
 RESULT 36
 ABP51669
 ID ABP51669 standard; peptide; 14 AA.
 XX AC ABP51669;
 XX DT 01-OCT-2002 (first entry)
 XX DE Thrombopoietin (TPO) agonist mimetic peptide SEQ ID NO:1.
 XX KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;
 KW complementarity determining region; immunoglobulin; antianaemic;
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis.
 XX KW
 XX OS Homo sapiens.

OS Synthetic.
 XX PN WO200246238-A2.
 XX PN 13-JUN-2002.
 XX PD
 XX PF
 XX PR 05-DEC-2001; 2001WO-US047656.
 XX PR 05-DEC-2000; 2000US-0251448P.
 XX PR 04-MAY-2001; 2001US-0288889P.
 XX PR 29-MAY-2001; 2001US-0294068P.
 XX PR (ALEX-) ALEXION PHARM INC.
 XX PA
 XX PI Bowdish KS, Barbas-Frederickson S, Renshaw M;
 XX PI WPI; 2002-566610/60.
 XX DR
 XX PT A novel immunogen molecule comprising a region in which amino acid
 PT residues corresponding to at least a portion of the complementary
 PT determining region are replaced or fused with an erythropoietin or
 PT thrombopoietin mimetic.
 XX PT
 XX PS Claim 18; Page 6; 113pp; English.
 XX CC The present invention describes an immunoglobulin molecule or its fragment
 CC (I) comprising a region where amino acid residues corresponding to at
 CC least a portion of the complementary determining region (CDR) are
 CC replaced or fused with biologically active peptides e.g. a peptide
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,
 CC that is flanked with proline at its carboxy terminus. (I) has
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as
 CC a stimulator of proliferation, differentiation and maturation of
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful
 CC for stimulating proliferation, differentiation or growth of
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with
 CC promegakaryocytes or megakaryocytes, which results in increased platelet
 CC production. (I) with a region where amino acid residues corresponding to
 CC a portion of CDR is replaced with an EPO mimetic, or which has one or
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the
 CC production of red blood cells, where (I) is contacted with haematopoietic
 CC stem cells or their progenitors. (I) is useful for diagnostics or
 CC therapeutics, in cell isolation strategies, and for treating patients
 CC suffering from deficiency in cell populations caused by disease,
 CC disorders or treatments related to the suppression of haematopoiesis.
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used in
 CC the exemplification of the present invention
 XX CC
 XX SQ Sequence 14 AA;
 Query Match 100.0%; Score 56; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGPTRLQWL A 10
 DB |||||
 2 EGPTRLQWL A 11
 RESULT 37
 AAE18011
 ID AAE18011 standard; peptide; 14 AA.
 XX AC AAE18011;
 XX DT 07-MAY-2002 (first entry)
 XX DE Human ligand #31 attached to an adenoviral vector.
 XX KW Human; adenoviral coat protein; non-native ligand; cell-surface receptor;
 KW therapy; anti-tumour agent; tumour necrosis factor; cancer; brain; lung;
 KW ovary; breast; prostate.
 XX KW

```

OS Homo sapiens.
XX
XX PN WO200192549-A2.
XX
XX PD 06-DEC-2001.
XX
XX PP 30-MAY-2001; 2001WO-US017391.
XX
XX PR 31-MAY-2000; 2000US-0208451P.
XX
XX PR 02-AUG-2000; 2000US-00631191.
XX
XX PA (GENV-) GENVEC INC.
XX
XX PI Wickham TJ, Kovesdi I, Roelvink PW, Einfeld D, Brough DE;
XX Lizonova A;
XX
XX DR WPI; 2002-147620/19.
XX
XX PT Adenoviral coat protein which permits production of adenoviral vectors
XX that bind and infect host cells not naturally infected by adenovirus,
XX comprises various non-native ligands.
XX
XX PS Claim 4; Page 45; 45pp; English.
XX
XX CC The invention relates to adenoviral coat proteins comprising various non-
XX native ligands. The invention provides a method of controlled gene
XX expression utilising selectively replicating competence and also a method
XX and a composition for targeting an adenoviral vector. A system
XX comprising a cell having a non-native cell-surface receptor, and a virus
XX having a non-native ligand which binds the non-native cell-surface
XX receptor of the cell is useful for propagating a virus and also for
XX assaying gene function. The system is also useful for isolating a nucleic
XX acid encoding a product comprising a desired property. Further the system
XX is useful for identifying functionally related coding sequences.
XX
XX CC Adenoviral vector comprising a non-native nucleic acid encoding a
XX therapeutic agent such as anti-tumour agent, preferably tumour necrosis
XX factor and a second non-native nucleic acid encoding an agent that
XX facilitates imaging and a targeting agent is useful for treating an
XX animal. The therapeutic agent can be used to treat cancer of the brain,
XX lung, ovary, breast and prostate. The present sequence is human non-
XX native ligand attached to an adenoviral vector
XX
XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11

RESULT 38
ABG71747
ID ABG71747 standard; peptide; 14 AA.
XX
XX AC ABG71747;
XX
XX DT 20-JAN-2003 (first entry)
XX
XX DE TPO receptor, MPL, agonist peptide, AF12505.
XX
XX KW Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
XX complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
XX thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
XX T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
XX proliferation; growth; differentiation; haematopoietic cell;
XX platelet progenitor cell; immune disorder; thrombocytopenia;
XX disseminated intravascular coagulation; stem cell; transplantation;
XX gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant.
XX
XX OS Synthetic.

XX
XX PN WO200278612-A2.
XX
XX PD 10-OCT-2002.
XX
XX PP 02-APR-2002; 2002WO-US010301.
XX
XX PR 02-APR-2001; 2001US-0281183P.
XX
XX PA (PURD ) PURDUE PHARMA LP.
XX
XX PI Soltis DA, Burch RM, Ogert RA;
XX
XX DR WPI; 2003-040615/03.
XX
XX PT New thrombopoietin synthebodies, useful for stimulating proliferation,
XX growth, or differentiation of hematopoietic cells, for treating or
XX preventing hematopoietic or immune disorders, e.g. thrombocytopenia.
XX
XX PS Claim 62; Page 71; 97pp; English.
XX
XX CC The invention discloses a variant of an immunoglobulin (Ig) variable
XX heavy or light chain domain that comprises at least one complementarity
XX determining region (CDR) and framework regions flanking the CDR. The CDR
XX also has added or substituted to it, at least one binding sequence which
XX is heterologous to the CDR and is an antigenic, agonistic sequence from a
XX thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
XX can be a binding sequence heterologous to the CDR, a cytotoxic T-
XX lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
XX cell sequence or a combination of each. The variant or thrombopoietin
XX synthebody, pharmaceutical and vaccine compositions are useful for
XX stimulating proliferation, growth or differentiation of haematopoietic
XX cells, particularly platelet progenitor cells. The variants are also
XX useful for treating or preventing haematopoietic or immune disorders
XX resulting from chemotherapy, radiation therapy, or bone marrow
XX transfusions (e.g. thrombocytopenia or disseminated intravascular
XX coagulation). Compositions comprising the synthebodies can be used for
XX the mobilisation, amplification and ex vivo expansion of stem cells and
XX committed precursor cells for autologous and allogeneic transplantation
XX as well as for the expansion of stem cells for gene therapy. They are
XX also useful as diagnostic or analytical reagents for studying the
XX function of thrombopoietin and its receptor in vivo or in vitro. The
XX sequence presented is the TPO receptor (MPL) agonist peptide, AF12505
XX
XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db |||||
2 EGPTLRQWLA 11

RESULT 39
ABR62907
ID ABR62907 standard; peptide; 14 AA.
XX
XX AC ABR62907;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE Thrombopoietin mimetic peptide AF12505.
XX
XX KW Thrombopoietin; mimetic; thrombocytopenia; antibody targeting.
XX
XX OS Synthetic.
XX
XX PN WO2003059251-A2.
XX
XX PD 24-JUL-2003.
XX

```

PF 22-OCT-2002; 2002WO-US033991.
XX
PR 22-OCT-2001; 2001US-0344614P.
PR 19-SEP-2002; 2002US-0412455P.
XX
XX (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Rader C, Sinha SC, Lerner R;
XX
DR WPI; 2003-636673/60.
XX
XX Antibody targeting compound useful e.g. for diagnostic immunoassays and
PT treating microbial diseases comprises targeting or biological agent
PT covalently linked to combining site of the antibody.
XX
XX Example 7; Page 62; 56pp; English.
XX
XX The present sequence is that of thrombopoietin (TPO) mimetic peptide
CC AF12505, which mimics the activity of recombinant TPO. The invention
CC provides antibody targeting compounds that are used to reprogram the
CC specificity of an antibody. The antibody targeting compound is linked to
CC the combining site of the antibody, such that the modified antibody takes
CC on the binding specificity of the targeting agent. In an example from the
CC invention, a TPO receptor targeting antibody compound was prepared by
CC covalently linking peptide AB12505 to aldolase monoclonal antibody 38C2.
CC The TPO receptor targeting antibody compound can be used to treat
CC thrombocytopaenia resulting from chemotherapy and bone marrow
CC transplantation
XX
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
RESULT 40
ADC33697
ID ADC33697 standard; peptide; 14 AA.
XX
AC ADC33697;
XX
DT 18-DEC-2003 (first entry)
XX
DE Erythropoietin receptor/erythropoietin consensus peptide SEQ ID NO:1.
XX
KW chimeric retrovirus envelope protein; ecotropic envelope protein;
KW cytosstatic; gene therapy; cancer.
XX
OS Synthetic.
XX
PN WO2003076596-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US007323.
XX
PR 08-MAR-2002; 2002US-0362655P.
XX
XX (UYMA-) UNIV MASSACHUSETTS.
XX
XX Green MR, Gollan TJ;
PI
XX WPI; 2003-722332/68.
XX
XX New chimeric retrovirus envelope protein comprising an ecotropic envelope
PT protein and a heterologous short peptide ligand inserted within the
PT ecotropic envelope protein useful for treating cancer.
XX

PS Disclosure; SEQ ID NO 1; 42pp; English.
XX
CC The present invention describes a chimeric retrovirus envelope protein
CC (1) comprising an ecotropic envelope protein and a heterologous short
CC peptide ligand inserted within the ecotropic envelope protein. Also
CC described: (1) a nucleic acid molecule comprising a sequence encoding the
CC recombinant chimeric envelope protein; (2) a vector comprising a nucleic
CC acid sequence encoding the chimeric envelope protein; (3) a recombinant
CC retroviral particle comprising a chimeric envelope protein comprising a
CC heterologous short peptide ligand; (3) altering retroviral tropism; (4)
CC identifying a nucleic acid sequence encoding the chimeric envelope
CC protein that alters viral tropism; (5) delivering a nucleic acid sequence
CC to a cell; and (6) treating cancer. (1) has cytostatic activity and can
CC be used in gene therapy. The chimeric retrovirus envelope protein is
CC useful for treating cancer, which comprises providing a cancer cell, e.g.
CC human cancer cell and infecting the cancer cell with a virus, e.g.
CC retrovirus comprising the chimeric envelope protein comprising a
CC heterologous short peptide ligand and a therapeutically useful gene, e.g.
CC encoding thymidine kinase. The present sequence represents an
CC erythropoietin receptor/erythropoietin consensus peptide, which is given
CC in the exemplification of the present invention.
XX
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 56; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
RESULT 41
ADJ73004
ID ADJ73004 standard; peptide; 14 AA.
XX
AC ADJ73004;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 458.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Scallion BJ, Ghayab J;
PI
XX WPI; 2003-804237/75.
XX
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
XX Disclosure; SEQ ID NO 458; 97pp; English.
XX
XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody

CC fragment where a protein has been inserted into, or replaces a portion
 CC of, one or more CDR regions, such that each CDR mimetibody comprises at
 CC least one portion of a heavy chain or light chain variable region, which
 CC itself comprises at least one human framework region and at least one
 CC ligand binding region (LBR). The present invention describes human
 CC mimetibodies, including modified immunoglobulins and cleavage products
 CC that can be useful in gene therapy and the generation of transgenic
 CC plants and animals. Furthermore, the CDR mimetibody is useful for
 CC preparing compositions for modulating, treating or reducing the symptoms
 CC of immune, cardiovascular, infectious, malignant and/or neurologic
 CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
 CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
 CC peptide sequence is a TPO mimetic peptide sequence used to make a
 CC mimetibody of the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
 |||||
 Db 2 EGPTLRQWLA 11

RESULT 42

ADN59652
 ID ADN59652 standard; peptide; 14 AA.

AC ADN59652;

DT 01-JUL-2004 (first entry)

XX Thrombopoietin mimetic peptide (TMP), seq id 1.

XX Haemostatic; antianaemic; immunosuppressive; platelet;
 XX transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
 XX TMP; c-mpl receptor; platelet precursor; megakaryocyte;
 XX thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;
 XX autoimmune haemolytic anaemia; Hughes's syndrome;
 XX lupoid thrombocytopaenia.

XX Homo sapiens.

XX WO2003031589-A2.

XX 17-APR-2003.

XX 11-OCT-2002; 2002WO-US032552.

XX 11-OCT-2001; 2001US-0328666P.

XX 10-OCT-2002; 2002US-00269806.

XX (AMGE-) AMGEN INC.

XX Min H, Sitney KC, Hartley C;

XX WPI; 2003-403101/38.

XX Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
 XX which stimulate the production of platelets and/or the production of
 XX platelet precursors, useful for treating thrombocytopaenia.

XX Disclosure; SEQ ID NO 1; 126pp; English.

XX The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
 CC binds to the c-mpl (mpl) receptor, and which stimulates the production of
 CC platelets and/or the production of platelet precursors, is new. Further
 CC disclosed is a composition of matter (II) that binds to an mpl receptor,
 CC and a pharmaceutical composition comprising (II) and a carrier. The
 CC pharmaceutical composition of the invention is useful for treating
 CC thrombocytopaenia in an animal, and for increasing megakaryocytes or

CC platelets in a patient. The TMP of the invention is useful for treating
 CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.
 CC disease conditions involving thrombocytopaenia such as aplastic anaemia,
 CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,
 CC autoimmune haemolytic anaemia, Hughes's syndrome and lupoid
 CC thrombocytopaenia. The TMP of the invention is also useful for
 CC maintaining the viability or storage life of platelets and/or
 CC megakaryocytes and its derived cells. The compounds demonstrate an
 CC improved ability to bind to and/or trigger transmembrane signal through,
 CC i.e. activating, the mpl receptor the compounds have superior
 CC thrombopoietic activity, i.e. the ability to stimulate, in vivo and in
 CC vitro, the production of platelets and/or megakaryocytocytic activity,
 CC i.e. the ability to stimulate, in vivo and in vitro, the production of,
 CC platelet precursors. Further, certain of the compounds also exhibit
 CC superior therapeutic properties, such as improved plasma half-life,
 CC biological activity and in vivo circulation time. The current sequence
 CC represents a TMP of the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 7; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
 |||||
 Db 2 EGPTLRQWLA 11

RESULT 43

ADJ52639

ID ADJ52639 standard; peptide; 14 AA.

XX AC ADJ52639;

XX 06-MAY-2004 (first entry)

XX CHI deleted mimetibody-related peptide SeqID458.

XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
 XX hypotensive; neuroprotective; nootropic; antibacterial; virucide;
 XX fungicide; gene therapy; immune disorder; cardiovascular disease;
 XX arrhythmia; hypertension; heart failure; neurodegenerative;
 XX multiple sclerosis; dementia; Alzheimer's disease; anaemia;
 XX cancerous condition; infectious disease; bacterial infection;
 XX viral infection; fungal infection.

XX Unidentified.

XX Synthetic.

XX WO2004002417-A2.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US020347.

XX 28-JUN-2002; 2002US-0392431P.

XX (CENZ) CENTOCOR INC.

XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;

XX Kutoloski KA;

XX WPI; 2004-082870/08.

XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
 XX modulating, treating, alleviating, preventing an immune, cardiovascular,
 XX or neurodegenerative disease or disorder, anemia, cancer, or infectious
 XX diseases.

XX Claim 2; SEQ ID NO 458; 129pp; English.

XX This invention relates to CHI deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an immunosuppressive,
 CC cardiovascular, cardiac, hypotensive, neuroprotective, nootropic,
 CC antibacterial, virucide or fungicide activity. In addition, the disclosed
 CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
 CC is useful for diagnosing or treating a disease condition in a cell,
 CC tissue, organ or animal, specifically for modulating, treating,
 CC alleviating, preventing the incidence or reducing the symptoms of an
 CC immune, cardiovascular (for example arrhythmia, hypertension or heart
 CC failure), or neurodegenerative (for example multiple sclerosis, dementia
 CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
 CC conditions, or infectious diseases (for example bacterial, viral or
 CC fungal infection). The present sequence is that of a peptide which may be
 CC used during the creation of a mimetibody of the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTRLQWLA 10
 |||||
 Db 2 EGPTRLQWLA 11

RESULT 44
 ADJ51600
 ID ADJ51600 standard; peptide; 14 AA.

XX AC ADJ51600;

XX DT 06-MAY-2004 (first entry)

XX DE CHI deleted mimetibody-related peptide SeqID458.

XX KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
 KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
 KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
 KW anti-allergic; muscular-Gen; cystostatic; anti-inflammatory; neuroleptic;
 KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
 KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
 KW dental disorder; oral disorder; dermatological disorder; ear disorder;
 KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
 KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
 KW obstetric disorder; haematologic disorder; immunological disorder;
 KW allergic disorder; infectious disorder; musculoskeletal disorder;
 KW oncological disorder; neurological disorder; nutritional disorder;
 KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
 KW renal disorder; pulmonary disorder.

XX OS Unidentified.

XX OS Synthetic.

XX FN WO2004002424-A2.

XX PD 08-JAN-2004.

XX PF 30-JUN-2003; 2003WO-US020495.

XX PR 28-JUN-2002; 2002US-0392431P.

XX PR 19-SEP-2002; 2002US-0412144P.

XX PA (CENZ) CENTOCOR INC.

XX XX Heavenr GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;

XX PI Kutoloski KA;

XX XX WPI; 2004-082872/08.

XX XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
 PT diagnosing, preventing or treating cardiovascular, dermatologic, and
 PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX PS Claim 14; SEQ ID NO 458; 123pp; English.

XX CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
 CC which encode them), compositions, methods and uses. The invention may be
 CC useful for the development of compounds with an osteopathic,
 CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
 CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
 CC immunomodulator, anti-allergic, muscular-Gen, cytostatic,
 CC anti-inflammatory, neuroleptic, ophthalmological, nephrotropic or
 CC respiratory-Gen activity acting as a tumour necrosis factor (TNF) -
 CC present invention are useful for the diagnosis, prevention and/or
 CC treatment of diseases or conditions associated with aberrant expression
 CC or activity of the CHI deleted mimetibody, such as a bone or joint,
 CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
 CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
 CC obstetric, haematologic, immunological, allergic, infectious,
 CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
 CC pediatric, psychiatric, renal or pulmonary disorders. The present
 CC sequence is that of a peptide which may be used during the creation of a
 CC mimetibody of the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTRLQWLA 10
 |||||
 Db 2 EGPTRLQWLA 11

RESULT 45

ADL27293
 ID ADL27293 standard; protein; 14 AA.

XX AC ADL27293;

XX DT 03-JUN-2004 (first entry)

XX DE Amino acid sequence of a thrombopoietin agonist peptide.

XX KW fusion protein; C4bp; alpha chain; systemic lupus erythematosus.

XX OS Homo sapiens.

XX OS Synthetic.

XX FN WO2004020639-A2.

XX PD 11-MAR-2004.

XX PF 12-AUG-2003; 2003WO-EP008928.

XX PR 14-AUG-2002; 2002EP-00292043.

XX XX (AVID-) AVIDIS SA.

XX XX Garnier L, Hill F, Julien M;

XX XX WPI; 2004-239202/22.

XX XX Obtaining a recombinant fusion protein, useful for treating lupus,
 PT comprises providing a prokaryotic host cell carrying a nucleic acid
 PT encoding the recombinant protein operably linked to a promoter functional
 PT in the prokaryotic cell.

XX PS Claim 8; Page 48; 69pp; English.

XX CC The specification describes a method for obtaining a recombinant fusion
 CC protein comprising a scaffold of a C-terminal core protein of C4bp alpha

chain, where the recombinant fusion protein is capable of forming multimers in soluble form in a prokaryotic host cell. The method comprises providing a prokaryotic host cell carrying a nucleic acid encoding the recombinant protein operably linked to a promoter functional in the prokaryotic cell, culturing the host cell under conditions where the recombinant protein is expressed, and recovering the recombinant protein where the protein is recovered in multimeric form without performing a scaffold refolding step. The protein is useful for treating systemic lupus erythematosus. The present sequence represents a thrombopoietin agonist peptide, which is used to produce fusion proteins of the invention.

XX Sequence 14 AA;
SQ

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 46
ADM72503
ID ADM72503 standard; peptide; 14 AA.
XX
AC ADM72503;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KM stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.

Key Location/Qualifiers
FT Modified-site 14 /note="sarcosine"
FT
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
XX Increasing hematopoietic stem cell production in subject, useful in reducing the incidence of delayed primary engraftment, comprises administering a Thrombopoietin mimetic compound e.g., a peptide to a subject.

XX Disclosure; Fig 2; 32pp; English.

XX The invention relates to a method (M1) for increasing haematopoietic stem cell production in a subject which involves administering a Thrombopoietin (TPO) mimetic compound to the subject. Also included is another method (M2) of providing haematopoietic stem cells to a subject which involves administering a TPO mimetic compound to a subject to enhance expansion of a stem cell population within bone marrow and/or mobilize stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or the stem cells in the peripheral circulation, and transplanting the harvested stem cells into the subject.

CC A method (M3) is also provided for reducing a time to engraftment following reinfusion of stem cells in a subject, involves administering a TPO mimetic compound to the subject, enhancing the expansion of the stem cell population within bone marrow and/or mobilizing the stem cells in peripheral circulation, harvesting one or more of the bone marrow stem cells or one or more of the stem cells in the peripheral circulation, and transplanting the one or more harvested stem cells into the subject. TPO mimetic compounds are disclosed as peptides, including cyclic or modified peptides. (M1) is useful for increasing haematopoietic stem cell production in a subject e.g., human. (M3) is useful for reducing time to engraftment following reinfusion of stem cells, reducing the incidence of delayed primary engraftment, reducing the time of platelet and/or neutrophil engraftment following reinfusion of stem cells in a subject. (M1) is also useful for increasing the number of stem cells from a donor whose cells are then used for rescue of recipient subject. Also useful in the treatment of thrombocytopenia. (M1) enables transplantation to proceed in patients who would not otherwise be considered as candidates because of unacceptably high risk of failed engraftment, reduces the number of aphereses required to generate a minimum acceptable harvest, reduces the incidence of primary and secondary failure of engraftment by increasing the number of haematopoietic stem cells (HSCs) available for transplantation and reduces the time required for primary engraftment. The present sequence represents an example of TPO mimetic peptide fragment.

XX Sequence 14 AA;
SQ

Query Match 100.0%; Score 56; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db 2 EGPTLRQWLA 11
|||||

RESULT 47
ADM72483
ID ADM72483 standard; peptide; 14 AA.
XX
AC ADM72483;
XX
DT 17-JUN-2004 (first entry)
XX
DE TPO mimetic peptide fragment.
XX
KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
KM stem cell therapy; HSC; transplantation; engraftment; mimetic.
XX
OS Synthetic.
XX
PN WO2004026332-A1.
XX
PD 01-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US029701.
XX
PR 18-SEP-2002; 2002US-0411700P.
PR 18-SEP-2002; 2002US-0411779P.
XX
PA (THRE-) 3-DIMENSIONAL PHARM INC.
XX
PI Kaushansky K, Macdonald BR;
XX
DR WPI; 2004-283153/26.
XX
XX Increasing hematopoietic stem cell production in subject, useful in reducing the incidence of delayed primary engraftment, comprises administering a Thrombopoietin mimetic compound e.g., a peptide to a subject.

XX Disclosure; Fig 2; 32pp; English.

XX The invention relates to a method (M1) for increasing haematopoietic stem
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing haematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, enhancing the expansion of the stem
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells in a subject.
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.

XX Sequence 14 AA;
 CC
 CC Query Match 100.0%; Score 56; DB 8; Length 14;
 CC Best Local Similarity 100.0%; Pred. No. 0.006;
 CC Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGGTLROWLA 10
 Db |||||
 2 EGGTLROWLA 11

RESULT 48
 ADM72495
 ID ADM72495 standard; peptide; 14 AA.
 XX
 AC ADM72495;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE TPO mimetic peptide fragment.
 XX
 KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 13
 FT /note= "p-amino-Phe"
 FT
 XX WO2004026332-A1.
 XX
 XX 01-APR-2004.
 XX
 XX 18-SEP-2003; 2003WO-US029701.
 XX
 XX 18-SEP-2002; 2002US-0411700P.

PR 18-SEP-2002; 2002US-0411779P.
 XX (THRE-) 3-DIMENSIONAL PHARM INC.
 PA Kaushansky K, Macdonald BR;
 XX WPT; 2004-283153/26.
 XX Increasing hematopoietic stem cell production in subject, useful in
 PT reducing the incidence of delayed primary engraftment, comprises
 PT administering a Thrombopoietin mimetic compound e.g., a peptide to a
 PT subject.
 XX
 PS Disclosure; Fig 2; 32pp; English.
 XX
 CC The invention relates to a method (M1) for increasing haematopoietic stem
 CC cell production in a subject which involves administering a
 CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is
 CC another method (M2) of providing haematopoietic stem cells to a subject
 CC which involves administering a TPO mimetic compound to a subject to
 CC enhance expansion of a stem cell population within bone marrow and/or
 CC mobilize stem cells in peripheral circulation, harvesting one or more of
 CC the bone marrow stem cells or the stem cells in the peripheral
 CC circulation, and transplanting the harvested stem cells into the subject.
 CC A method (M3) is also provided for reducing a time to engraftment
 CC following reinfusion of stem cells in a subject, involves administering a
 CC TPO mimetic compound to the subject, enhancing the expansion of the stem
 CC cell population within bone marrow and/or mobilizing the stem cells in
 CC peripheral circulation, harvesting one or more of the bone marrow stem
 CC cells or one or more of the stem cells in the peripheral circulation, and
 CC transplanting the one or more harvested stem cells into the subject. TPO
 CC mimetic compounds are disclosed as peptides, including cyclic or modified
 CC peptides. (M1) is useful for increasing haematopoietic stem cell
 CC production in a subject e.g., human. (M3) is useful for reducing time to
 CC engraftment following reinfusion of stem cells, reducing the incidence of
 CC delayed primary engraftment, reducing the time of platelet and/or
 CC neutrophil engraftment following reinfusion of stem cells in a subject.
 CC (M1) is also useful for increasing the number of stem cells in a subject.
 CC whose cells are then used for rescue of recipient subject. Also useful in
 CC the treatment of thrombocytopenia. (M1) enables transplantation to
 CC proceed in patients who would not otherwise be considered as candidates
 CC because of unacceptably high risk of failed engraftment, reduces the
 CC number of aphereses required to generate a minimum acceptable harvest,
 CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of TPO mimetic peptide
 CC fragment.

XX Sequence 14 AA;
 CC
 CC Query Match 100.0%; Score 56; DB 8; Length 14;
 CC Best Local Similarity 100.0%; Pred. No. 0.006;
 CC Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGGTLROWLA 10
 Db |||||
 2 EGGTLROWLA 11

RESULT 49
 ADM72487
 ID ADM72487 standard; peptide; 14 AA.
 XX
 AC ADM72487;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE TPO mimetic peptide fragment.
 XX
 KW TPO; haematopoietic stem cell; thrombopoietin; haemostatic;
 KW stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 14

XX FT /label= bala

XX FT /note= "beta-alanine"

XX PN WO2004026332-A1.

XX XX 01-APR-2004.

XX PD 18-SEP-2003; 2003WO-US029701.

XX PF 18-SEP-2002; 2002US-0411700P.

XX PR 18-SEP-2002; 2002US-0411779P.

XX XX (THRE-) 3-DIMENSIONAL PHARM INC.

XX PA Kaushansky K, Macdonald BR;

XX PI WPI; 2004-283153/26.

XX DR Increasing hematopoietic stem cell production in subject, useful in

XX PT reducing the incidence of delayed primary engraftment, comprises

XX FT administering a Thrombopoietin mimetic compound e.g., a peptide to a

XX FT subject.

XX XX Disclosure; Fig 2; 32pp; English.

XX PS The invention relates to a method (M1) for increasing haematopoietic stem

XX CC cell production in a subject which involves administering a

XX CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is

XX CC another method (M2) of providing haematopoietic stem cells to a subject

XX CC which involves administering a TPO mimetic compound to a subject to

XX CC enhance expansion of a stem cell population within bone marrow and/or

XX CC mobilize stem cells in peripheral circulation, harvesting one or more of

XX CC the bone marrow stem cells or the stem cells in the peripheral

XX CC circulation, and transplanting the harvested stem cells into the subject.

XX CC A method (M3) is also provided for reducing a time to engraftment

XX CC following reinfusion of stem cells in a subject, involves administering a

XX CC TPO mimetic compound to the subject, enhancing the expansion of the stem

XX CC cell population within bone marrow and/or mobilizing the stem cells in

XX CC peripheral circulation, harvesting one or more of the bone marrow stem

XX CC cells or one or more of the stem cells in the peripheral circulation, and

XX CC transplanting the one or more harvested stem cells into the subject. TPO

XX CC mimetic compounds are disclosed as peptides, including cyclic or modified

XX CC peptides. (M1) is useful for increasing haematopoietic stem cell

XX CC production in a subject e.g., human. (M3) is useful for reducing time to

XX CC engraftment following reinfusion of stem cells, reducing the incidence of

XX CC delayed primary engraftment, reducing the time of platelet and/or

XX CC of platelet production and reducing the time of stem cells in a subject.

XX CC (M1) is also useful for increasing the number of stem cells from a donor

XX CC whose cells are then used for rescue of recipient subject. Also useful in

XX CC the treatment of thrombocytopenia. (M1) enables transplantation to

XX CC proceed in patients who would not otherwise be considered as candidates

XX CC because of unacceptably high risk of failed engraftment, reduces the

XX CC number of aphereses required to generate a minimum acceptable harvest,

XX CC reduces the incidence of primary and secondary failure of engraftment by

XX CC increasing the number of haematopoietic stem cells (HSCs) available for

XX CC transplantation and reduces the time required for primary engraftment.

XX CC The present sequence represents an example of TPO mimetic peptide

XX CC fragment.

XX SQ Sequence 14 AA;

Query March 100.0%; Score 56; DB 8; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGPTLRQWLA 10

|||||||

Db 2 EGPTLRQWLA 11

RESULT 50

ADM72497

ID ADM72497 standard; peptide; 14 AA.

XX AC ADM72497;

XX DT 17-JUN-2004 (first entry)

XX DE TPO mimetic peptide fragment.

XX KW TPO; haematopoietic stem cell; thrombopoietin; haemoetatic;

XX OS stem cell therapy; HSC; transplantation; engraftment; mimetic.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 13

XX FT /note= "Ac-Lys"

XX PN WO2004026332-A1.

XX PD 01-APR-2004.

XX PF 18-SEP-2003; 2003WO-US029701.

XX PR 18-SEP-2002; 2002US-0411700P.

XX PR 18-SEP-2002; 2002US-0411779P.

XX XX (THRE-) 3-DIMENSIONAL PHARM INC.

XX PA Kaushansky K, Macdonald BR;

XX PI WPI; 2004-283153/26.

XX DR Increasing hematopoietic stem cell production in subject, useful in

XX PT reducing the incidence of delayed primary engraftment, comprises

XX FT administering a Thrombopoietin mimetic compound e.g., a peptide to a

XX FT subject.

XX XX Disclosure; Fig 2; 32pp; English.

XX PS The invention relates to a method (M1) for increasing haematopoietic stem

XX CC cell production in a subject which involves administering a

XX CC Thrombopoietin (TPO) mimetic compound to the subject. Also included is

XX CC another method (M2) of providing haematopoietic stem cells to a subject

XX CC which involves administering a TPO mimetic compound to a subject to

XX CC enhance expansion of a stem cell population within bone marrow and/or

XX CC mobilize stem cells in peripheral circulation, harvesting one or more of

XX CC the bone marrow stem cells or the stem cells in the peripheral

XX CC circulation, and transplanting the harvested stem cells into the subject.

XX CC A method (M3) is also provided for reducing a time to engraftment

XX CC following reinfusion of stem cells in a subject, involves administering a

XX CC TPO mimetic compound to the subject, enhancing the expansion of the stem

XX CC cell population within bone marrow and/or mobilizing the stem cells in

XX CC peripheral circulation, harvesting one or more of the bone marrow stem

XX CC cells or one or more of the stem cells in the peripheral circulation, and

XX CC transplanting the one or more harvested stem cells into the subject. TPO

XX CC mimetic compounds are disclosed as peptides, including cyclic or modified

XX CC peptides. (M1) is useful for increasing haematopoietic stem cell

XX CC production in a subject e.g., human. (M3) is useful for reducing time to

XX CC engraftment following reinfusion of stem cells, reducing the incidence of

XX CC delayed primary engraftment, reducing the time of platelet and/or

XX CC of platelet production and reducing the time of stem cells in a subject.

XX CC (M1) is also useful for increasing the number of stem cells from a donor

XX CC whose cells are then used for rescue of recipient subject. Also useful in

XX CC the treatment of thrombocytopenia. (M1) enables transplantation to

XX CC proceed in patients who would not otherwise be considered as candidates

XX CC because of unacceptably high risk of failed engraftment, reduces the

XX CC number of aphereses required to generate a minimum acceptable harvest,

XX CC reduces the incidence of primary and secondary failure of engraftment by

XX CC increasing the number of haematopoietic stem cells (HSCs) available for

XX CC transplantation and reduces the time required for primary engraftment.

XX CC The present sequence represents an example of TPO mimetic peptide

XX CC fragment.

XX SQ Sequence 14 AA;

Query March 100.0%; Score 56; DB 8; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EGPTLRQWLA 10

|||||||

CC reduces the incidence of primary and secondary failure of engraftment by
 CC increasing the number of haematopoietic stem cells (HSCs) available for
 CC transplantation and reduces the time required for primary engraftment.
 CC The present sequence represents an example of IPO mimetic peptide
 CC fragment.

XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 56; DB 8; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
 |||||
 Db 2 EGPTLRQWLA 11

Search completed: May 12, 2006, 10:36:47
 Job time : 77.7496 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 12.1368 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-62
Perfect score: 56
Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*

1: Pirl:1*
2: Pirl:2*
3: Pirl:3*
4: Pirl:4*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	75.0	1019	2	Tl1560
2	40	71.4	296	2	AG0147
3	39	69.6	131	2	S74539
4	39	69.6	217	2	S46354
5	39	69.6	473	2	E84853
6	39	69.6	656	2	S30484
7	39	69.6	656	2	S30483
8	39	69.6	1034	1	GNLJCA
9	39	69.6	1035	1	GNLJGG
10	39	69.6	1036	1	GNLJG2
11	39	69.6	1055	1	GNLJST
12	39	69.6	1055	2	S53092
13	38	67.9	134	2	B75468
14	38	67.9	1712	1	CGHU2B
15	37.5	67.0	333	2	A36925
16	37	66.1	306	2	D70601
17	37	66.1	306	2	T45453
18	37	66.1	547	2	T36550
19	37	66.1	777	2	A87309
20	37	66.1	816	2	A71006
21	37	66.1	825	2	JC4163
22	37	66.1	1039	2	S46347
23	37	66.1	1058	2	S08436
24	37	66.1	1083	2	S54293
25	36	64.3	151	2	S63748
26	36	64.3	200	2	T23485
27	36	64.3	207	2	T37464
28	36	64.3	278	2	T45494
29	36	64.3	312	2	F86876

30	36	64.3	370	2	AB2578
31	36	64.3	370	2	A97360
32	36	64.3	481	2	T49321
33	36	64.3	530	2	A81958
34	36	64.3	531	2	E81015
35	36	64.3	827	2	S29955
36	36	64.3	855	1	JQ2003
37	36	64.3	855	1	JQ2004
38	36	64.3	855	2	F45557
39	36	64.3	856	1	A45394
40	36	64.3	856	1	VCLJPP
41	36	64.3	856	2	S23823
42	36	64.3	857	2	S19886
43	36	64.3	907	2	AB1885
44	36	64.3	1032	2	S12153
45	36	64.3	1056	1	GNLJG3
46	36	64.3	1123	2	T51517
47	36	64.3	1257	2	S28764
48	36	64.3	1339	2	T40245
49	35.5	63.4	602	2	T45278
50	35	62.5	120	2	A97655
51	35	62.5	120	2	AG2878
52	35	62.5	197	2	G82973
53	35	62.5	297	2	B87109
54	35	62.5	415	2	I39527
55	35	62.5	451	2	A95379
56	35	62.5	472	1	WZEDC
57	35	62.5	472	2	C85571
58	35	62.5	472	2	E30720
59	35	62.5	480	2	H84747
60	35	62.5	486	2	B86411
61	35	62.5	501	2	S48120
62	35	62.5	505	2	T19971
63	35	62.5	506	2	T19973
64	35	62.5	521	2	T01923
65	35	62.5	524	2	E71881
66	35	62.5	527	2	B84633
67	35	62.5	610	2	A86176
68	35	62.5	612	2	S71221
69	35	62.5	635	2	A87433
70	35	62.5	691	2	A54741
71	35	62.5	721	2	A39707
72	35	62.5	791	2	A82291
73	35	62.5	793	2	JC5831
74	35	62.5	795	2	H70934
75	35	62.5	803	2	E82392
76	35	62.5	904	2	C70559
77	35	62.5	1040	2	T08190
78	35	62.5	1091	2	G59435
79	35	62.5	1380	2	T25284
80	35	62.5	3345	2	T13423
81	34	60.7	70	2	T06920
82	34	60.7	85	2	F75291
83	34	60.7	98	2	A70301
84	34	60.7	158	2	A10605
85	34	60.7	158	2	E64823
86	34	60.7	158	2	E90745
87	34	60.7	184	2	H83409
88	34	60.7	220	2	AC0318
89	34	60.7	267	2	I40327
90	34	60.7	295	2	S76136
91	34	60.7	302	2	E75414
92	34	60.7	315	2	G95415
93	34	60.7	339	2	A83358
94	34	60.7	350	2	H85953
95	34	60.7	350	2	H85953
96	34	60.7	350	2	E91108
97	34	60.7	379	2	I48133
98	34	60.7	379	2	I48132
99	34	60.7	379	2	I48180
100	34	60.7	379	2	I48134
101	34	60.7	382	2	T34789
102	34	60.7	413	2	JC6512

membrane-bound lyt
outer membrane lip
related to heterok
probable permease
ABC transporter, p
surface glycoprote
env polypyrprotein -
env polypyrprotein -
external glycoprot
env polypyrprotein pr
env polypyrprotein pr
env polypyrprotein -
hypothetical prote
HIV-1 retropepsin
telomerase reverse
neurocan precursor
probable transcrip
oligopeptide ABC t
hypothetical prote
conserved hypotet
transcription regu
integrase/recombin
lambda integrase f
HYPOHETICAL 50.8
deoxyribodipyrimid
hypothetical prote
deoxyribodipyrimid
probable steroid d
protein F3M18.4 li
deoxyribodipyrimid
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable deoxyribo
hypothetical prote
erythrocyte membra
c-di-GMP phosphodi
kinesin-related pr
hypothetical prote
phosphoenolpyruvat
probable polA prot
hypothetical prote
DLC-1 (deleted in
hypothetical prote
hypothetical prote
ribosomal protein
hypothetical prote
ribosomal protein
conserved hypotet
hypothetical prote
probable sensory t
hypothetical prote
probable nicotinat
baf protein - Bord
hypothetical prote
PecN-related prote
probable LysR-fam
hypothetical prote
A/G-specific adeni
adenine glycosylas
adenine glycosylas
ubiquinol-cytochro
ubiquinol-cytochro
ubiquinol-cytochro
ubiquinol-cytochro
hypothetical prote
methyltransferase

103	34	60.7	458	2	B83372	hypothetical prote	176	33	58.9	219	2	S32081	RNA-directed DNA p
104	34	60.7	463	2	S27491	hypothetical prote	177	33	58.9	219	2	S32139	RNA-directed DNA p
105	34	60.7	469	2	AD1926	hypothetical prote	178	33	58.9	219	2	S32078	RNA-directed DNA p
106	34	60.7	470	2	AD0888	Sufi protein [impo	179	33	58.9	219	2	S32066	RNA-directed DNA p
107	34	60.7	470	2	E91116	suppressor of ftsI	180	33	58.9	219	2	S32126	RNA-directed DNA p
108	34	60.7	470	2	B85961	suppressor of ftsI	181	33	58.9	219	2	S32074	RNA-directed DNA p
109	34	60.7	470	2	G65088	sufi protein precu	182	33	58.9	219	2	S32053	RNA-directed DNA p
110	34	60.7	571	2	AI0506	probable sulfatase	183	33	58.9	219	2	S32088	RNA-directed DNA p
111	34	60.7	594	2	G83878	L-lactate permease	184	33	58.9	219	2	S32072	RNA-directed DNA p
112	34	60.7	750	2	AE7501	topoisomerase iv c	185	33	58.9	219	2	S32070	RNA-directed DNA p
113	34	60.7	750	2	AE7519	topoisomerase iv s	186	33	58.9	219	2	S32087	RNA-directed DNA p
114	34	60.7	868	2	C81200	aconitate hydratase	187	33	58.9	219	2	S32076	RNA-directed DNA p
115	34	60.7	868	2	B81775	aconitate hydratase	188	33	58.9	219	2	S32092	RNA-directed DNA p
116	34	60.7	955	2	T48515	hypothetical prote	189	33	58.9	219	2	S32160	RNA-directed DNA p
117	34	60.7	1009	2	S28081	pol polyprotein -	190	33	58.9	219	2	S32054	RNA-directed DNA p
118	34	60.7	1054	1	GNLJG5	HIV-1 retropepsin	191	33	58.9	219	2	S32135	RNA-directed DNA p
119	34	60.7	1061	1	GNLJG4	HIV-1 retropepsin	192	33	58.9	219	2	S32159	RNA-directed DNA p
120	34	60.7	1123	2	T28139	PK4 protein kinase	193	33	58.9	219	2	S32069	RNA-directed DNA p
121	34	60.7	1155	2	AC2426	adenylate cyclase	194	33	58.9	219	2	S32063	RNA-directed DNA p
122	34	60.7	1168	2	T30935	reverse transcript	195	33	58.9	219	2	S32128	RNA-directed DNA p
123	34	60.7	1707	2	A33526	collagen alpha 2(I	196	33	58.9	219	2	S32085	RNA-directed DNA p
124	34	60.7	2329	2	S44625	CSOC3.6 protein -	197	33	58.9	219	2	S32064	RNA-directed DNA p
125	34	60.7	2363	2	T38841	probable pre-mRNA	198	33	58.9	219	2	S32056	RNA-directed DNA p
126	34	60.7	3433	1	GNWVKV	genome polyprotein	199	33	58.9	219	2	S32059	RNA-directed DNA p
127	33.5	59.8	209	2	B87627	hypothetical prote	200	33	58.9	219	2	S32131	RNA-directed DNA p
128	33	58.9	39	2	I37555	homeobox - human (201	33	58.9	219	2	S32095	RNA-directed DNA p
129	33	58.9	75	2	S03065	ser protein - huma	202	33	58.9	219	2	S32071	RNA-directed DNA p
130	33	58.9	85	2	AH1010	conserved hypothet	203	33	58.9	219	2	S32127	RNA-directed DNA p
131	33	58.9	129	2	S63743	HIV-1 retropepsin	204	33	58.9	219	2	S32134	RNA-directed DNA p
132	33	58.9	130	2	S63754	HIV-1 retropepsin	205	33	58.9	219	2	S32093	RNA-directed DNA p
133	33	58.9	132	2	S63733	HIV-1 retropepsin	206	33	58.9	219	2	S32075	RNA-directed DNA p
134	33	58.9	136	2	S63746	HIV-1 retropepsin	207	33	58.9	219	2	S32079	RNA-directed DNA p
135	33	58.9	140	2	S63737	HIV-1 retropepsin	208	33	58.9	219	2	S32119	RNA-directed DNA p
136	33	58.9	141	2	S63750	HIV-1 retropepsin	209	33	58.9	219	2	S32138	RNA-directed DNA p
137	33	58.9	142	2	AF0961	heat shock protein	210	33	58.9	219	2	S32129	RNA-directed DNA p
138	33	58.9	143	2	S63735	HIV-1 retropepsin	211	33	58.9	219	2	S32097	RNA-directed DNA p
139	33	58.9	145	2	S63745	HIV-1 retropepsin	212	33	58.9	219	2	S32117	RNA-directed DNA p
140	33	58.9	149	2	S63734	HIV-1 retropepsin	213	33	58.9	219	2	S32077	RNA-directed DNA p
141	33	58.9	151	2	S63731	HIV-1 retropepsin	214	33	58.9	219	2	S32098	RNA-directed DNA p
142	33	58.9	153	2	A96751	hypothetical prote	215	33	58.9	219	2	S32051	RNA-directed DNA p
143	33	58.9	154	2	AC0496	heat shock protein	216	33	58.9	219	2	A46097	GPI-anchor biosynt
144	33	58.9	160	2	S63753	HIV-1 retropepsin	217	33	58.9	219	2	S32132	RNA-directed DNA p
145	33	58.9	162	2	S63732	HIV-1 retropepsin	218	33	58.9	219	2	S32058	RNA-directed DNA p
146	33	58.9	170	2	B83871	hypothetical prote	219	33	58.9	219	2	S32140	RNA-directed DNA p
147	33	58.9	176	2	S63747	HIV-1 retropepsin	220	33	58.9	224	2	F70614	probable malonyl c
148	33	58.9	190	2	AD2343	conserved hypothet	221	33	58.9	235	2	C83822	hypothetical prote
149	33	58.9	208	2	AD2808	conserved hypothet	222	33	58.9	260	2	E81194	biotin synthesis p
150	33	58.9	213	2	AC3389	uracil-DNA glycosy	223	33	58.9	262	2	I40321	divIB protein - Ba
151	33	58.9	213	2	AI2233	hypothetical prote	224	33	58.9	264	2	AG2095	hypothetical prote
152	33	58.9	214	1	ASLJG5	vif protein - simi	225	33	58.9	302	2	S75227	hypothetical prote
153	33	58.9	214	1	ASLJSM	vif protein - simi	226	33	58.9	311	1	RGECK	regulatory protein
154	33	58.9	217	2	B97587	hypothetical prote	227	33	58.9	311	2	AH0867	transcription acti
155	33	58.9	219	2	S32096	RNA-directed DNA p	228	33	58.9	311	2	C85936	positive regulator
156	33	58.9	219	2	S32157	RNA-directed DNA p	229	33	58.9	311	2	H91090	positive regulator
157	33	58.9	219	2	S32080	RNA-directed DNA p	230	33	58.9	313	2	T35826	probable dehydrog
158	33	58.9	219	2	S32062	RNA-directed DNA p	231	33	58.9	327	2	B82277	probable transpos
159	33	58.9	219	2	S32047	RNA-directed DNA p	232	33	58.9	331	2	E48445	glyceraldehyde-3-p
160	33	58.9	219	2	S32065	RNA-directed DNA p	233	33	58.9	336	2	F72389	hypothetical prote
161	33	58.9	219	2	S32122	RNA-directed DNA p	234	33	58.9	337	2	T27635	homeobox protein c
162	33	58.9	219	2	S32060	RNA-directed DNA p	235	33	58.9	338	2	AH3321	transcription regu
163	33	58.9	219	2	S32118	RNA-directed DNA p	236	33	58.9	341	2	AG0195	probable exported
164	33	58.9	219	2	S32057	RNA-directed DNA p	237	33	58.9	352	2	C83636	conserved hypothet
165	33	58.9	219	2	S32048	RNA-directed DNA p	238	33	58.9	364	2	C84187	hypothetical prote
166	33	58.9	219	2	S32094	RNA-directed DNA p	239	33	58.9	380	2	A71390	ubiquitin-cyclocho
167	33	58.9	219	2	S32061	RNA-directed DNA p	240	33	58.9	383	2	T20572	hypothetical prote
168	33	58.9	219	2	S32084	RNA-directed DNA p	241	33	58.9	389	2	B69096	corrinoid/iron-eul
169	33	58.9	219	2	S32152	RNA-directed DNA p	242	33	58.9	410	2	H86290	hypothetical prote
170	33	58.9	219	2	S32089	RNA-directed DNA p	243	33	58.9	416	2	B69374	N-ethylmelamine ch
171	33	58.9	219	2	S32073	RNA-directed DNA p	244	33	58.9	431	2	A81150	histidyl-tRNA synt
172	33	58.9	219	2	S32049	RNA-directed DNA p	245	33	58.9	451	2	AF3053	conserved hypothet
173	33	58.9	219	2	S32120	RNA-directed DNA p	246	33	58.9	451	2	G98232	hypothetical prote
174	33	58.9	219	2	S32086	RNA-directed DNA p	247	33	58.9	472	2	A13367	multidrug resistan
175	33	58.9	219	2	S32133	RNA-directed DNA p	248	33	58.9	477	2	H83588	probable MPS trans

249	33	58.9	499	2	H83254	probable MPS trans	322	57.1	218	2	JC7220	nuclear protein SR
250	33	58.9	510	2	S41943	cellulose 1,4-beta	323	57.1	220	2	D82139	conserved hypothet
251	33	58.9	511	2	S44716	cellulose 1,4-beta	324	57.1	221	2	B36881	MPM2-reactive phos
252	33	58.9	536	1	SYECBB	2,3-dihydroxybenzo	325	57.1	222	2	B90888	hypothetical prote
253	33	58.9	536	2	E85558	2,3-dihydroxybenzo	326	57.1	229	2	JC7219	nuclear protein SR
254	33	58.9	536	2	A99708	2,3-dihydroxybenzo	327	57.1	244	2	A82316	hypothetical prote
255	33	58.9	540	2	S41942	cellulose 1,4-beta	328	57.1	249	2	E87575	ABC transporter. A
256	33	58.9	542	1	A54963	transcription fact	329	57.1	252	2	C84522	HAC transporter. A
257	33	58.9	544	2	A72459	probable glutamyl-	330	57.1	254	2	T35994	22 kDa peroxisomal
258	33	58.9	559	2	B47175	reverse transcript	331	57.1	255	2	A45881	probable ABC-type
259	33	58.9	559	2	A47175	reverse transcript	332	57.1	269	2	T35999	MHC class II hista
260	33	58.9	567	2	T16105	hypothetical prote	333	57.1	274	2	S05509	hypothetical prote
261	33	58.9	582	2	T16104	hypothetical prote	334	57.1	278	2	F84127	hypothetical prote
262	33	58.9	586	2	B84271	glutamyl-tRNA synt	335	57.1	285	2	G71337	probable dimethyla
263	33	58.9	589	2	F87626	sensor histidine k	336	57.1	288	2	S73016	H+-transporting tw
264	33	58.9	600	2	C83221	transport protein	337	57.1	291	2	T10966	6-phosphogluconate
265	33	58.9	618	2	A10171	hypothetical integ	338	57.1	293	2	T35157	translocation prot
266	33	58.9	625	2	T40742	hypothetical prote	339	57.1	302	2	T36030	probable fructokin
267	33	58.9	629	2	T16767	NADH2 dehydrogenas	340	57.1	309	2	F83434	L-fucose-specific
268	33	58.9	633	2	B70946	comE protein - Sym	341	57.1	310	2	JC7853	hypothetical prote
269	33	58.9	709	2	S75212	conserved hypothet	342	57.1	312	2	S76507	probable UV damage
270	33	58.9	719	2	B95325	poliphosphate kina	343	57.1	326	2	C75350	hypothetical prote
271	33	58.9	739	2	T45429	probable ppk prote	344	57.1	327	2	H82736	hypothetical prote
272	33	58.9	742	2	E70673	transcription-rela	345	57.1	331	2	A72534	hypothetical prote
273	33	58.9	772	2	E82592	probable ATP-depen	346	57.1	335	2	T20825	hypothetical prote
274	33	58.9	838	2	B83150	pol polyprotein -	347	57.1	362	2	G65119	protein serine/thr
275	33	58.9	902	2	T01668	pol polyprotein -	348	57.1	368	2	G65119	hypothetical 40.4
276	33	58.9	912	2	S33980	alanine-tRNA ligas	349	57.1	373	2	B95291	conserved hypothet
277	33	58.9	924	2	E71476	DNA-directed DNA p	350	57.1	393	2	B85992	probable transport
278	33	58.9	994	2	AC3480	hypothetical prote	351	57.1	394	2	C82439	peptide methionine
279	33	58.9	1000	2	T30280	HIV-1 retropepsin	352	57.1	396	1	XNECD	aspartate transami
280	33	58.9	1002	1	GNLJND	pol polyprotein -	353	57.1	396	2	AD0616	aspartate aminotra
281	33	58.9	1002	1	S54378	HIV-1 retropepsin	354	57.1	396	2	A85619	aspartate aminotra
282	33	58.9	1003	1	GNVWLV	HIV-1 retropepsin	355	57.1	396	2	C90755	probable transport
283	33	58.9	1003	1	B44001	HIV-1 retropepsin	356	57.1	399	2	F91146	probable malate de
284	33	58.9	1003	1	GNVWA2	HIV-1 retropepsin	357	57.1	400	2	T18570	hypothetical membr
285	33	58.9	1003	2	T09440	pol polyprotein -	358	57.1	408	2	T50876	thiamine pyridinyl
286	33	58.9	1012	1	GNVWVL	HIV-1 retropepsin	359	57.1	409	2	T47118	aspartate transami
287	33	58.9	1015	1	GNVWH3	HIV-1 retropepsin	360	57.1	418	2	S46315	interferon consens
288	33	58.9	1027	1	GNLJSI	recombination acti	361	57.1	424	2	A35861	hypothetical prote
289	33	58.9	1073	2	I51055	probable arabinosy	362	57.1	425	2	A45064	membrane protein,
290	33	58.9	1094	2	F70697	HIV-1 retropepsin	363	57.1	438	2	G87337	glutamate-cysteine
291	33	58.9	1124	1	GNLJFP	HIV-1 retropepsin	364	57.1	438	2	G87337	aspartate transami
292	33	58.9	1124	2	B45557	pol polyprotein -	365	57.1	454	2	T06136	deoxyribodipyrimid
293	33	58.9	1124	2	S23820	major core protein	366	57.1	456	2	T06136	deoxyribodipyrimid
294	33	58.9	1233	1	P3XR03	probable polyketid	367	57.1	473	2	A10587	hypothetical prote
295	33	58.9	1733	2	D70887	polyketide synthas	368	57.1	475	2	T01352	glutamyl-tRNA(Gln)
296	33	58.9	1784	2	E86921	toxin-like outer m	369	57.1	482	2	D75346	multidrug-efflux t
297	33	58.9	2299	2	H71879	toxin-like outer m	370	57.1	489	2	A11276	multidrug-efflux t
298	33	58.9	2529	2	B64635	genome polyprotein	371	57.1	489	2	A11639	drug-export protei
299	33	58.9	3430	1	GNVWVW	CCHH finger protei	372	57.1	493	2	AD1398	occludin - chicken
300	32.5	58.0	235	2	S55883	hypothetical prote	373	57.1	493	2	AG1773	hypothetical prote
301	32.5	58.0	245	2	A84279	probable transcript	374	57.1	504	2	A49467	conserved hypothet
302	32.5	58.0	480	2	T36822	hypothetical prote	375	57.1	519	2	AD2315	hypothetical prote
303	32	57.1	84	2	AB2301	conserved hypothet	376	57.1	519	2	D82536	glutamate synthase
304	32	57.1	111	2	H95908	hypothetical prote	377	57.1	525	2	D90048	multidrug resistanc
305	32	57.1	117	2	AB2757	hypothetical prote	378	57.1	532	1	E69821	adult-specific 61.
306	32	57.1	123	2	T49336	holo-(acyl-carrier	379	57.1	532	1	E69821	DNA repair protein
307	32	57.1	125	2	F81197	hypothetical prote	380	57.1	559	2	A10135	hypothetical prote
308	32	57.1	147	2	T38472	hypothetical prote	381	57.1	560	2	S65148	tRNA adenyltrans
309	32	57.1	183	2	AC0767	dTDP-4-dehydrotham	382	57.1	568	2	T05218	hypothetical prote
310	32	57.1	183	2	S15302	hypothetical prote	383	57.1	573	2	I60247	SEC23 protein homo
311	32	57.1	183	2	S15302	probable carbonic	384	57.1	575	2	A11996	hypothetical prote
312	32	57.1	193	2	T35847	probable phosphopa	385	57.1	594	2	B82994	probable secretion
313	32	57.1	195	2	F86017	hypothetical 21.8K	386	57.1	601	2	T36323	flagellar hook-ass
314	32	57.1	195	2	S47694	hypothetical prote	387	57.1	606	2	G64659	penicillin-binding
315	32	57.1	196	2	A10502	molybdopterin bios	388	57.1	610	2	AE2436	probable exported
316	32	57.1	198	2	B84382	hypothetical prote	389	57.1	615	2	A10615	probable amidase
317	32	57.1	207	2	B75327	hypothetical prote	390	57.1	615	2	H90754	probable amidase y
318	32	57.1	209	2	A83849	alpha-ribazole-5'-	391	57.1	615	2	F85618	ycbB protein precu
319	32	57.1	210	2	G85729	hypothetical prote	392	57.1	615	2	F85618	
320	32	57.1	214	2	S07989	vif protein - simi	393	57.1	615	2	F85618	
321	32	57.1	214	2	T11561	vif protein - simi	394	57.1	615	2	D64832	

395	32	57.1	623	2	T48859	disease resistance	468	31	55.4	178	2	AB2743	hypothetical prote
396	32	57.1	623	2	T06674	hypothetical prote	469	31	55.4	189	2	F95017	hypothetical prote
397	32	57.1	629	2	T06675	hypothetical prote	470	31	55.4	189	2	G97890	hypothetical prote
398	32	57.1	664	2	G89894	protein kinase [im	471	31	55.4	203	2	H95842	probable glutathio
399	32	57.1	691	2	T44543	probable bacterioph	472	31	55.4	209	2	D95274	hypothetical prote
400	32	57.1	715	2	S38051	DOA1 protein - Yea	473	31	55.4	216	2	D75567	GTP cyclohydrolase
401	32	57.1	725	2	A11544	conserved hypotet	474	31	55.4	218	2	B71183	hypothetical prote
402	32	57.1	728	2	D86278	hypothetical prote	475	31	55.4	219	2	E75143	phosphoglycolate p
403	32	57.1	755	2	D75598	photoreceptor - De	476	31	55.4	219	2	B97461	structure of gluta
404	32	57.1	761	2	T09052	hypothetical prote	477	31	55.4	219	2	AD2679	Glutathione-S-tran
405	32	57.1	763	2	E96571	hypothetical prote	478	31	55.4	219	2	T38159	probable RNA-bind
406	32	57.1	765	2	T09574	transport protein	479	31	55.4	233	2	A83862	initiation of chro
407	32	57.1	775	2	B64319	carbon-monoxide de	480	31	55.4	247	2	S18604	triose-phosphate i
408	32	57.1	798	2	C98069	primosomal replica	481	31	55.4	249	2	A70840	hypothetical prote
409	32	57.1	808	2	B64914	dimethylsulfoxide	482	31	55.4	260	2	F81830	conserved hypotet
410	32	57.1	808	2	E90915	probable oxidoredu	483	31	55.4	260	2	C64129	biotin synthetase p
411	32	57.1	808	2	B85764	probable oxidoredu	484	31	55.4	261	2	G84010	hypothetical prote
412	32	57.1	916	2	H72372	exonuclease ABC c	485	31	55.4	262	2	B83126	probable transcrip
413	32	57.1	916	2	G75417	SNF2/Rad54 helicase	486	31	55.4	263	2	A30227	hypothetical prote
414	32	57.1	921	2	A40537	DNA-directed DNA p	487	31	55.4	264	2	AH0759	propanediol diffu
415	32	57.1	942	2	S75598	poly(A) polymerase	488	31	55.4	264	2	AH2202	hypothetical prote
416	32	57.1	943	2	T34847	probable transcrip	489	31	55.4	264	2	G83343	hypothetical prote
417	32	57.1	1003	2	T13856	ker protein - frui	490	31	55.4	264	2	H69838	chloride peroxidase
418	32	57.1	1016	1	S40838	formate dehydrogen	491	31	55.4	268	2	B42424	chitinase (EC 3.2.
419	32	57.1	1016	2	D91231	formate dehydrogen	492	31	55.4	268	2	C84314	hypothetical prote
420	32	57.1	1019	2	T00117	dye protein - frui	493	31	55.4	272	2	T35231	hypothetical prote
421	32	57.1	1099	1	S31926	myosin IB heavy ch	494	31	55.4	281	2	H86159	hypothetical prote
422	32	57.1	1099	2	A59300	myosin-If - mouse	495	31	55.4	286	2	G70537	hypothetical prote
423	32	57.1	1102	2	S55100	hypothetical prote	496	31	55.4	289	2	JQ0059	hypothetical 31.6K
424	32	57.1	1170	2	S03308	cell surface glyco	497	31	55.4	291	2	F69959	hypothetical prote
425	32	57.1	1175	2	T46124	hypothetical prote	498	31	55.4	292	2	F86431	hypothetical prote
426	32	57.1	1268	2	S52781	neurocan - mouse	499	31	55.4	294	2	S18542	hypothetical prote
427	32	57.1	1333	2	S30356	CDC25 protein homo	500	31	55.4	295	2	E75366	glutamyl-tRNA synt
428	32	57.1	1369	2	T03104	tegument protein h	501	31	55.4	307	2	B83167	hypothetical prote
429	32	57.1	1541	2	S46686	hypothetical prote	502	31	55.4	326	2	C24430	glycerolaldehyde-3-p
430	32	57.1	1545	2	T26589	hypothetical prote	503	31	55.4	329	2	S55489	probable galactosy
431	32	57.1	1616	2	T17884	S-layer protein -	504	31	55.4	329	2	S73710	adhesin pl homolog
432	32	57.1	1620	2	H83281	conserved hypotet	505	31	55.4	332	2	T04754	hypothetical prote
433	32	57.1	2108	2	H70819	probable polyketid	506	31	55.4	335	2	H70545	probable fahB prot
434	31.5	56.2	332	2	F83600	conserved hypotet	507	31	55.4	336	1	DEP2G	glycerolaldehyde-3-p
435	31.5	56.2	909	2	D87434	phosphoenolpyruvat	508	31	55.4	337	2	S45080	glycerolaldehyde-3-p
436	31.5	56.2	911	2	AD0959	Two-component sens	509	31	55.4	338	1	DEIS3C	glycerolaldehyde-3-p
437	31.5	56.2	1080	2	T03964	probable ubiquitin	510	31	55.4	338	2	JQ1287	glycerolaldehyde-3-p
438	31	55.4	77	2	A83468	hypothetical prote	511	31	55.4	338	2	E90998	hypothetical prote
439	31	55.4	85	2	B91147	hypothetical prote	512	31	55.4	339	2	AG0799	probable transcrip
440	31	55.4	85	2	F85932	hypothetical prote	513	31	55.4	342	1	S64042	porphobilinogen sy
441	31	55.4	85	2	C65120	hypothetical 10.0	514	31	55.4	346	2	D85818	unknown protein en
442	31	55.4	93	2	B70967	hypothetical prote	515	31	55.4	355	2	AG1877	hypothetical prote
443	31	55.4	108	2	G84522	similar to gibbere	516	31	55.4	359	2	D83103	probable phospholi
444	31	55.4	108	2	T49731	hypothetical prote	517	31	55.4	382	2	T11138	ubiquinol-cytochro
445	31	55.4	110	2	G90584	SOS ribosomal prot	518	31	55.4	382	2	S77130	hypothetical prote
446	31	55.4	111	1	RSYM22	ribosomal protein	519	31	55.4	387	2	AI0511	probable metabolit
447	31	55.4	111	2	G70521	hypothetical prote	520	31	55.4	392	2	T51151	probable nuclear D
448	31	55.4	112	2	S75152	hypothetical prote	521	31	55.4	396	2	S63374	diphosphomevalonat
449	31	55.4	113	2	S10612	ribosomal protein	522	31	55.4	400	2	C87021	serine-threonine p
450	31	55.4	114	2	H69636	ribosomal protein	523	31	55.4	406	2	AB1060	arginine desminase
451	31	55.4	114	2	AI0294	probable membrane	524	31	55.4	407	2	A86298	hypothetical prote
452	31	55.4	117	2	H90021	SOS ribosomal prot	525	31	55.4	407	2	T36404	probable monooxyge
453	31	55.4	118	2	AC1403	ribosomal protein	526	31	55.4	409	2	T37507	aspartate transami
454	31	55.4	118	2	AB1779	ribosomal protein	527	31	55.4	412	2	S76130	hypothetical prote
455	31	55.4	125	2	B82378	hypothetical prote	528	31	55.4	416	1	JC4952	transcription init
456	31	55.4	144	2	B91207	heat shock protein	529	31	55.4	417	2	F97789	ampG protein [limp
457	31	55.4	144	2	G65170	heat shock protein	530	31	55.4	422	2	F96826	hypothetical prote
458	31	55.4	144	2	D86053	heat shock protein	531	31	55.4	424	2	T25774	hypothetical prote
459	31	55.4	147	2	A70928	hypothetical prote	532	31	55.4	425	2	A87517	aminotransferase,
460	31	55.4	150	2	S37014	transposase (clone	533	31	55.4	425	2	S30558	ribose-phosphate d
461	31	55.4	153	2	A97524	hypothetical prote	534	31	55.4	427	2	B83001	probable C4-dicarb
462	31	55.4	153	2	C69193	ribosomal protein	535	31	55.4	430	2	F87472	PNX oxidoreductase
463	31	55.4	155	2	D75147	leu ribosomal prot	536	31	55.4	433	2	S51837	glycerolaldehyde-3-p
464	31	55.4	155	2	B71187	probable ribosomal	537	31	55.4	433	2	S51836	glycerolaldehyde-3-p
465	31	55.4	162	2	A34429	dihydrofolate redu	538	31	55.4	437	1	B29336	ubiquinol-cytochro
466	31	55.4	165	2	G69224	conserved hypotet	539	31	55.4	449	2	T48511	aspartate transami
467	31	55.4	174	2	F82786	conserved hypotet	540	31	55.4	450	2	E98303	hypothetical 49.3K

541	31	55.4	450	2	AH2979	nitrotriacetate	614	31	55.4	1793	2	T47897	guanine nucleotide
542	31	55.4	454	2	JC7231	thermophilic desul	615	31	55.4	1972	2	S68176	TOG protein - huma
543	31	55.4	454	2	F87149	proble secreted p	616	31	55.4	2156	1	RRVUNE	genome polypeptin
544	31	55.4	475	2	F84261	probable transport	617	31	55.4	3036	2	T18995	hypothetical prote
545	31	55.4	477	2	B83476	probable transport	618	31	55.4	3172	2	S22012	erythronolide synt
546	31	55.4	483	2	D71439	probable Beta-Amyl	619	31	55.4	3178	2	S13595	6-deoxyerythronoli
547	31	55.4	491	2	G89860	hypothetical prote	620	31	55.4	4725	1	A44357	dynem heavy chain
548	31	55.4	496	2	G87546	acid-CoA ligase, p	621	30.5	54.5	427	1	VHVNVP	nucleoprotein - Pi
549	31	55.4	498	2	H85190	probable beta-amy	622	30.5	54.5	529	2	AI0986	probable membrane
550	31	55.4	499	2	H87460	hypothetical prote	623	30.5	54.5	748	2	I48744	semaphorin A - mou
551	31	55.4	499	2	S39113	alpha-N-arabinofur	624	30.5	54.5	943	2	C82559	isoleucyl-tRNA syn
552	31	55.4	503	2	E83490	probable MFS trans	625	30.5	54.5	6420	2	T30283	polyketide synthas
553	31	55.4	511	1	E70391	major facilitator	626	30	53.6	65	2	S59524	triose-phosphate i
554	31	55.4	524	2	S75852	hypothetical prote	627	30	53.6	69	2	AB2606	transcription repr
555	31	55.4	526	2	C83488	probable semialdeh	628	30	53.6	109	2	AB3026	conserved hypothet
556	31	55.4	529	2	H75078	ATP-dependent RNA	629	30	53.6	113	2	S26277	T-cell receptor be
557	31	55.4	541	2	F84187	phytoene dehydrog	630	30	53.6	113	2	I38318	T-cell receptor be
558	31	55.4	548	2	T52556	beta-amyase (EC 3	631	30	53.6	113	2	S26276	hypothetical prote
559	31	55.4	557	2	T44843	glucose-6-phosphat	632	30	53.6	116	2	E71171	hypothetical prote
560	31	55.4	565	1	C72550	methylmalonyl-CoA	633	30	53.6	116	2	T06099	hypothetical prote
561	31	55.4	572	2	T47219	amino acid transpo	634	30	53.6	119	2	G98258	hypothetical prote
562	31	55.4	575	2	C86398	protein T7N9.26 [i	635	30	53.6	119	2	C89582	protein C18A11.1 l
563	31	55.4	589	2	T34878	probable integral	636	30	53.6	123	2	C35352	hypothetical prote
564	31	55.4	610	1	S30234	transcription fact	637	30	53.6	123	2	T43588	hypothetical prote
565	31	55.4	612	2	A13493	oligopeptide-bindi	638	30	53.6	123	2	S15322	hypothetical prote
566	31	55.4	628	2	S35547	double-strand-brea	639	30	53.6	142	2	AC0928	probable regulator
567	31	55.4	645	2	G88130	protein F10G7.4 [i	640	30	53.6	144	2	S38391	T-cell receptor be
568	31	55.4	651	2	T25953	hypothetical prote	641	30	53.6	145	2	S07957	hypothetical prote
569	31	55.4	664	2	H83962	serine/threonine p	642	30	53.6	148	2	H84491	hypothetical prote
570	31	55.4	750	2	D86245	hypothetical prote	643	30	53.6	150	2	T41638	probable protein k
571	31	55.4	762	2	H83415	cis/trans isomeraz	644	30	53.6	151	2	AB0313	conserved hypothet
572	31	55.4	766	2	E70916	probable bisC prot	645	30	53.6	154	2	F64026	hypothetical prote
573	31	55.4	775	2	AB3402	topoisomerase IV c	646	30	53.6	158	2	A85596	probable sensory t
574	31	55.4	777	1	G69773	conserved hypothet	647	30	53.6	158	2	D72305	hypothetical prote
575	31	55.4	789	2	S28259	androgen-regulated	648	30	53.6	165	2	G97231	diverged Metallo-d
576	31	55.4	792	2	A84308	chloride channel [649	30	53.6	172	1	NKVG02	core protein p19 -
577	31	55.4	801	2	A37353	membrane protein 4	650	30	53.6	172	1	JC1153	hypothetical 19.7K
578	31	55.4	802	2	T45642	FtsH metalloprotei	651	30	53.6	175	2	C69137	conserved hypothet
579	31	55.4	812	2	A80680	probable dimethyl	652	30	53.6	177	2	C69137	hypothetical prote
580	31	55.4	815	1	PHCEGG	phosphorylase (EC	653	30	53.6	180	2	T09083	MutR/nudix family
581	31	55.4	815	2	B86009	glycogen phosphory	654	30	53.6	185	2	D82418	recombinase Xfa001
582	31	55.4	815	2	A98163	glycogen phosphory	655	30	53.6	188	2	C82863	conserved hypothet
583	31	55.4	815	2	A80996	glycogen phosphory	656	30	53.6	194	2	C87507	hypothetical prote
584	31	55.4	825	2	A48537	starch branching e	657	30	53.6	195	2	A12549	conserved hypothet
585	31	55.4	836	2	T33222	hypothetical prote	658	30	53.6	196	2	A98071	conserved hypothet
586	31	55.4	838	2	A32262	fatty-acid synthas	659	30	53.6	197	2	AE2063	hypothetical prote
587	31	55.4	843	2	D96495	unknown protein [i	660	30	53.6	201	2	G59096	hypothetical prote
588	31	55.4	853	2	AB2020	hypothetical prote	661	30	53.6	202	2	S56267	probable membrane
589	31	55.4	885	2	G91080	hypothetical prote	662	30	53.6	206	1	I40173	orotate phosphorib
590	31	55.4	886	1	JC5085	replication licens	663	30	53.6	206	2	S38626	glutathione transf
591	31	55.4	899	2	H85925	hypothetical prote	664	30	53.6	206	2	AD1898	urase accessory p
592	31	55.4	929	2	G72677	hypothetical prote	665	30	53.6	207	2	C36961	hypothetical prote
593	31	55.4	934	2	E64235	hypothetical prote	666	30	53.6	207	2	AD2350	hypothetical prote
594	31	55.4	952	2	S32954	hypothetical prote	667	30	53.6	208	2	E86014	probable enzyme yh
595	31	55.4	952	2	E86147	hypothetical prote	668	30	53.6	208	2	E91168	probable enzyme li
596	31	55.4	971	2	T24866	TlK6.4 protein - A	669	30	53.6	208	2	S47687	hypothetical prote
597	31	55.4	988	2	H71338	hypothetical prote	670	30	53.6	209	1	XURTWC	methylated-DNA-lpr
598	31	55.4	1005	2	AB0946	conserved hypothet	671	30	53.6	211	2	AB3178	glutathione S-tran
599	31	55.4	1016	2	AB0946	probable lipoprote	672	30	53.6	211	2	C82969	sarcosine oxidase
600	31	55.4	1041	2	T31437	integrin alpha cha	673	30	53.6	214	2	G87680	conserved hypothet
601	31	55.4	1045	2	S23570	pol polyprotein ho	674	30	53.6	214	2	T22892	hypothetical prote
602	31	55.4	1140	2	D88690	protein F4IH10.3 [675	30	53.6	214	2	T22896	hypothetical prote
603	31	55.4	1171	1	Q0KBFP	pyruvate (flavodox	676	30	53.6	215	2	T22895	hypothetical prote
604	31	55.4	1202	2	S55553	LAR-interacting pr	677	30	53.6	215	2	E96533	hypothetical prote
605	31	55.4	1220	2	AD0125	exodeoxyribonuclea	678	30	53.6	221	2	F81876	hypothetical prote
606	31	55.4	1288	2	ED0363	mitogen-activated	679	30	53.6	221	2	A82190	hypothetical prote
607	31	55.4	1298	1	EDBE75	immediate-early pr	680	30	53.6	224	2	H70717	hypothetical prote
608	31	55.4	1411	2	S55123	hypothetical prote	681	30	53.6	224	2	T36146	hypothetical prote
609	31	55.4	1429	2	S06434	homeotic protein l	682	30	53.6	224	2	B84059	hypothetical prote
610	31	55.4	1500	2	AF2027	hypothetical prote	683	30	53.6	227	2	T42015	probable RNA polym
611	31	55.4	1576	2	T28655	receptor-like hist	684	30	53.6	228	2	D64479	hypothetical prote
612	31	55.4	1576	2	T30898	receptor-like hist	685	30	53.6	230	2	T32999	hypothetical prote
613	31	55.4	1680	2	T01367	hypothetical prote	686	30	53.6	230	2	H86758	hypothetical prote

687	30	53.6	232	2	B82987	probable hydrolase	760	30	53.6	385	2	B83506	probable MPS trans
688	30	53.6	232	2	A60083	neural induction h	761	30	53.6	386	2	B82321	serine/threonine k
689	30	53.6	237	2	AC1953	rubrythrinn [impo	762	30	53.6	388	2	G70729	hypothetical prote
690	30	53.6	238	2	S44969	lmbv protein - Str	763	30	53.6	388	2	T28693	hypothetical prote
691	30	53.6	244	2	F82510	probable transcrip	764	30	53.6	389	2	S73910	probable serine/th
692	30	53.6	245	2	C82189	conserved hypotHet	765	30	53.6	391	2	D41880	isopropylmalate sy
693	30	53.6	245	2	AH0130	probable oxidoredu	766	30	53.6	393	2	T35894	hypothetical prote
694	30	53.6	252	2	B97072	probable hydrolase	767	30	53.6	394	2	AF2363	phosphomannose iso
695	30	53.6	256	2	G95866	probable pyroloquo	768	30	53.6	394	2	G98319	mannose-6-phosphat
696	30	53.6	260	2	S72767	probable ABC-type	769	30	53.6	395	2	S76793	hypothetical prote
697	30	53.6	266	2	B70872	probable ABC-type	770	30	53.6	403	2	H90827	probable head port
698	30	53.6	267	2	S55487	probable 3-methyl-	771	30	53.6	403	2	F85685	probable portal pr
699	30	53.6	276	2	B75337	hypothetical prote	772	30	53.6	405	2	T14311	aspartate transami
700	30	53.6	283	2	T31122	probable transcrip	773	30	53.6	406	2	A39339	protein C inhibito
701	30	53.6	290	2	D87281	conserved hypotHet	774	30	53.6	407	2	FC5124	aspartate transami
702	30	53.6	293	2	A82690	phosphatidylserine	775	30	53.6	407	2	F82323	arginine deiminase
703	30	53.6	295	2	AD0306	probable membrane	776	30	53.6	408	2	T31254	biphenyl dioxigena
704	30	53.6	297	2	S55063	conserved hypotHet	777	30	53.6	409	2	T51126	hypothetical prote
705	30	53.6	300	2	S75692	hypothetical prote	778	30	53.6	410	2	S02167	type I site-specif
706	30	53.6	303	2	AG0603	hypothetical ABC t	779	30	53.6	413	2	JV0087	glucose-1-phosphat
707	30	53.6	303	2	H64820	probable oligopept	780	30	53.6	413	2	F90773	periplasmic glucos
708	30	53.6	303	2	G90742	probable transport	781	30	53.6	413	2	B85636	periplasmic glucos
709	30	53.6	303	2	B85593	probable transport	782	30	53.6	414	1	S05441	cytochrome b5-rela
710	30	53.6	308	2	A82391	transcription regu	783	30	53.6	414	2	H97601	hypothetical prote
711	30	53.6	308	2	T36366	probable adenine g	784	30	53.6	414	2	AI2823	aminotransferase I
712	30	53.6	312	2	B69615	isochoriemataze (E	785	30	53.6	415	2	C87328	conserved hypotHet
713	30	53.6	316	2	D71271	hypothetical prote	786	30	53.6	417	1	S49051	cytochrome P450 ty
714	30	53.6	317	2	F93172	probable transcrip	787	30	53.6	419	2	T29266	hypothetical prote
715	30	53.6	317	2	T33887	hypothetical prote	788	30	53.6	419	2	H95919	probable reverse t
716	30	53.6	318	1	D23766	modulation protein	789	30	53.6	419	2	H96025	probable reverse t
717	30	53.6	321	2	C70653	probable prephenat	790	30	53.6	419	2	A95386	Reverse transcript
718	30	53.6	321	2	H87683	transcription regu	791	30	53.6	423	2	S41122	mannose-6-phosphat
719	30	53.6	322	2	F86918	probable prephenat	792	30	53.6	426	1	A48829	transcription fact
720	30	53.6	322	2	AD1535	hypothetical prote	793	30	53.6	430	2	S05481	keratin 18, type I
721	30	53.6	324	2	C82251	GGDEF family prote	794	30	53.6	435	2	B72418	conserved hypotHet
722	30	53.6	325	2	B82129	conserved hypotHet	795	30	53.6	435	2	S43507	translation elonga
723	30	53.6	332	2	B47017	probable transcrip	796	30	53.6	435	2	H90162	hypothetical prote
724	30	53.6	332	2	AD2541	transcription init	797	30	53.6	440	2	A83435	ATP synthase in ty
725	30	53.6	332	2	T44437	aminodeoxychromis	798	30	53.6	441	2	T32021	hypothetical prote
726	30	53.6	334	2	T29061	probable integral	799	30	53.6	445	1	C42022	transcription fact
727	30	53.6	335	1	A39862	protein-tyrosine-p	800	30	53.6	454	2	S16565	noli protein - Rhi
728	30	53.6	335	2	T17408	rRNA (adenine-N6-)	801	30	53.6	456	2	G64083	glmu protein homol
729	30	53.6	336	2	A64378	hypothetical prote	802	30	53.6	457	2	C82720	UDP-N-acetylglucos
730	30	53.6	336	2	T26189	hypothetical prote	803	30	53.6	463	2	G90857	hypothetical prote
731	30	53.6	338	2	C83570	hypothetical prote	804	30	53.6	463	2	G85508	hypothetical prote
732	30	53.6	340	2	AD0701	tetrathionate redu	805	30	53.6	469	2	B82527	multidrug efflux p
733	30	53.6	341	2	AG0358	conserved hypotHet	806	30	53.6	471	2	B38637	Ras inhibitor (clo
734	30	53.6	344	2	B84377	protein export [im	807	30	53.6	472	1	F69763	multidrug resistan
735	30	53.6	345	2	B85904	hypothetical prote	808	30	53.6	472	2	T20454	hypothetical prote
736	30	53.6	345	2	AI0830	probable RNA methy	809	30	53.6	473	2	G72205	hypothetical prote
737	30	53.6	345	2	G91059	hypothetical prote	810	30	53.6	475	2	G70861	hypothetical prote
738	30	53.6	345	2	D65036	hypothetical 37.8	811	30	53.6	476	2	H83265	probable D-alanyl-
739	30	53.6	347	2	AC3299	pseudouridyate sy	812	30	53.6	478	2	AG0170	conserved hypotHet
740	30	53.6	348	2	T51340	mitogen-activated	813	30	53.6	483	2	S74637	lysine decarboxyla
741	30	53.6	348	2	T52635	mitogen-activated	814	30	53.6	489	2	A45988	dentin matrix acid
742	30	53.6	350	2	C87536	transcription regu	815	30	53.6	492	2	H85783	probable aldehyde
743	30	53.6	350	2	AG0879	A/G-specific adeni	816	30	53.6	492	2	D90935	probable aldehyde
744	30	53.6	350	2	A40647	methylation-indepe	817	30	53.6	492	2	B64934	succinylglutamic s
745	30	53.6	350	2	T22701	hypothetical prote	818	30	53.6	498	2	JC5983	protein kinase C-i
746	30	53.6	351	2	S40754	glycerol-3-phospha	819	30	53.6	499	2	AC2068	cell death suppres
747	30	53.6	353	2	B72738	hypothetical prote	820	30	53.6	506	2	D90207	conserved hypotHet
748	30	53.6	353	2	AI0036	conserved hypotHet	821	30	53.6	507	2	S06272	photosystem II chl
749	30	53.6	354	2	T46683	probable lipopate p	822	30	53.6	508	1	QJ5B6A	photosystem II chl
750	30	53.6	355	2	C83354	hypothetical prote	823	30	53.6	508	1	QJLV6A	photosystem II chl
751	30	53.6	361	2	F91207	hypothetical prote	824	30	53.6	508	1	QJNT6A	photosystem II chl
752	30	53.6	361	2	H86053	hypothetical prote	825	30	53.6	508	1	QJRT6A	photosystem II chl
753	30	53.6	361	2	C55171	hypothetical 41.0	826	30	53.6	508	1	QJZMBB	photosystem II chl
754	30	53.6	366	2	B36919	hypothetical prote	827	30	53.6	508	2	S12132	photosystem II chl
755	30	53.6	371	2	D75266	cell division prot	828	30	53.6	508	2	S04100	photosystem II chl
756	30	53.6	371	2	A88520	41.8K hypotHetal	829	30	53.6	508	2	S01385	photosystem II chl
757	30	53.6	376	2	C87596	glycosyl hydrolase	830	30	53.6	508	2	S14140	photosystem II chl
758	30	53.6	378	2	S55634	glycoprotein M - e	831	30	53.6	508	2	S33774	photosystem II chl
759	30	53.6	378	2	D83381	hypothetical prote	832	30	53.6	508	2	S07546	photosystem II chl

833	30	53.6	508	2	T07985	906	30	53.6	813	2	G81694	phosphorylase (EC
834	30	53.6	508	2	T07277	907	30	53.6	814	2	A71540	phosphorylase (EC
835	30	53.6	508	2	T07508	908	30	53.6	814	2	T02041	1,4-alpha-glucan b
836	30	53.6	508	2	S12129	909	30	53.6	821	2	D64119	glycogen phosphory
837	30	53.6	509	2	JA0148	910	30	53.6	823	2	T06574	probable 1,4-alpha
838	30	53.6	509	2	AB1824	911	30	53.6	861	2	T36381	probable large ATP
839	30	53.6	509	2	T11996	912	30	53.6	862	2	T36380	probable large ATP
840	30	53.6	509	2	S73243	913	30	53.6	862	2	T29133	hypothetical prote
841	30	53.6	509	2	S78322	914	30	53.6	863	2	B97504	DNA polymerase [im
842	30	53.6	509	2	T06855	915	30	53.6	863	2	AE2722	RNA-directed RNA p
843	30	53.6	513	1	S50216	916	30	53.6	893	2	AH2007	toxin secretion AB
844	30	53.6	514	1	QJWMPB	917	30	53.6	897	2	T21688	hypothetical prote
845	30	53.6	514	2	B67378	918	30	53.6	906	2	G96621	probable disease r
846	30	53.6	515	2	G75267	919	30	53.6	908	2	B69435	signal-transducing
847	30	53.6	522	2	T44369	920	30	53.6	914	2	S18942	hypothetical prote
848	30	53.6	525	2	A29776	921	30	53.6	945	2	F84519	probable retroelem
849	30	53.6	534	2	A29776	922	30	53.6	955	2	S52959	male-specific leth
850	30	53.6	535	1	WMBE1	923	30	53.6	973	2	AB2340	hypothetical prote
851	30	53.6	535	2	C24187	924	30	53.6	984	2	T48216	hypothetical prote
852	30	53.6	535	2	S76564	925	30	53.6	1002	2	A36691	Ca2+-transporting
853	30	53.6	538	2	S76175	926	30	53.6	1008	2	S38003	translation elonga
854	30	53.6	544	2	T38469	927	30	53.6	1009	2	S44621	C50C3.2 protein -
855	30	53.6	551	2	D69282	928	30	53.6	1010	2	T36383	probable large ATP
856	30	53.6	551	2	D38640	929	30	53.6	1014	2	JE0333	kioto protein - r
857	30	53.6	553	2	T20192	930	30	53.6	1021	2	T00361	hypothetical prote
858	30	53.6	566	2	AC0559	931	30	53.6	1050	2	HW0092	serine-threonine k
859	30	53.6	566	2	T09154	932	30	53.6	1065	2	H95321	NOIG efflux transp
860	30	53.6	566	2	C90691	933	30	53.6	1075	2	D70568	hypothetical prote
861	30	53.6	566	2	T46219	934	30	53.6	1121	2	G64103	exodeoxyribonuclea
862	30	53.6	566	2	G95541	935	30	53.6	1139	2	T23018	hypothetical prote
863	30	53.6	566	2	G85541	936	30	53.6	1150	2	T40395	exodeoxyribonuclea
864	30	53.6	566	2	E64774	937	30	53.6	1191	2	AF2501	LIM domain protein
865	30	53.6	568	2	S57830	938	30	53.6	1197	2	D82696	WD-repeat protein
866	30	53.6	568	2	E90364	939	30	53.6	1200	2	T35539	hypothetical prote
867	30	53.6	569	2	S23542	940	30	53.6	1220	1	DJBEC3	probable DNA methy
868	30	53.6	569	2	S1806	941	30	53.6	1220	2	T42573	DNA-directed DNA P
869	30	53.6	569	2	S57831	942	30	53.6	1252	2	D71810	DNA-directed DNA P
870	30	53.6	571	2	T48737	943	30	53.6	1271	2	T43269	probable type II D
871	30	53.6	576	2	H70961	944	30	53.6	1279	2	E64709	microcystin synthet
872	30	53.6	583	2	AI3098	945	30	53.6	1435	1	BVBVL1	type IIS restricti
873	30	53.6	589	2	H84072	946	30	53.6	1558	2	T29253	guanine nucleotide
874	30	53.6	593	2	S45281	947	30	53.6	1583	2	S59644	hypothetical prote
875	30	53.6	606	2	C84390	948	30	53.6	1650	2	S28721	sister chromatid c
876	30	53.6	606	2	H98187	949	30	53.6	1985	2	I61776	hypothetical prote
877	30	53.6	623	2	D71435	950	30	53.6	2150	1	S27802	Munc13-2 - rat
878	30	53.6	634	2	T00054	951	30	53.6	2150	2	T19450	zinc finger protei
879	30	53.6	638	2	B83890	952	30	53.6	2157	2	AI3009	hypothetical prote
880	30	53.6	639	2	JO0607	953	30	53.6	2359	2	B96832	peptide synthetase
881	30	53.6	640	1	ALASGR	954	30	53.6	2420	2	A84652	hypothetical prote
882	30	53.6	640	2	A29166	955	30	53.6	2476	2	T34022	hypothetical prote
883	30	53.6	645	2	T29818	956	30	53.6	2566	2	E98274	zonadhesin - pig
884	30	53.6	648	2	T43337	957	30	53.6	3856	2	T51174	ataxia-telangiecta
885	30	53.6	652	2	T16582	958	30	53.6	5138	2	B96695	hypothetical prote
886	30	53.6	652	2	C07042	959	29.5	52.7	584	2	D84264	probable transcrip
887	30	53.6	656	2	AC0573	960	29.5	52.7	685	2	JC6331	hydrolase, alpha/b
888	30	53.6	676	2	H96970	961	29.5	52.7	686	2	E87490	methionine synthas
889	30	53.6	683	2	B71325	962	29.5	52.7	796	2	D97065	nirv precursor (AF
890	30	53.6	686	2	T08919	963	29.5	52.7	796	2	T00587	nitrite reductase,
891	30	53.6	693	2	AE0005	964	29.5	52.7	1080	2	G70513	hypothetical prote
892	30	53.6	696	2	T02832	965	29.5	52.7	1206	2	E87072	hypothetical prote
893	30	53.6	702	2	T04424	966	29.5	52.7	42	2	G86487	rho-type guanine e
894	30	53.6	716	2	G44490	967	29.5	52.7	63	2	T30614	NADH dehydrogenase
895	30	53.6	716	2	T51354	968	29.5	52.7	81	2	B81133	transketolase [imp
896	30	53.6	727	2	T08920	969	29.5	52.7	85	2	S76843	probable ubiquitin
897	30	53.6	729	2	T06797	970	29.5	52.7	87	2	S42171	5-methyltetrahydro
898	30	53.6	729	2	AD0485	971	29	51.8	87	2	T10462	hypothetical prote
899	30	53.6	762	2	H87466	972	29	51.8	92	2	PQ0632	hypothetical prote
900	30	53.6	780	2	D75361	973	29	51.8	92	2	PQ0632	coat protein - lil
901	30	53.6	781	2	T09358	974	29	51.8	92	2	PQ0632	coat protein - lil
902	30	53.6	789	2	F69099	975	29	51.8	92	2	PQ0632	coat protein - lil
903	30	53.6	795	2	T02981	976	29	51.8	92	2	PQ0632	coat protein - lil
904	30	53.6	799	2	T01663	977	29	51.8	92	2	PQ0632	coat protein - lil
905	30	53.6	804	2	H75125	978	29	51.8	92	2	PQ0632	coat protein - lil

979 29 51.8 92 2 PQ0628 coat protein - tul
980 29 51.8 94 1 D45390 trans-activating t
981 29 51.8 97 2 H42409 glutathione transf
982 29 51.8 104 2 E82797 conserved hypothet
983 29 51.8 108 1 RPECW trp operon repress
984 29 51.8 108 2 S45254 trp operon repress
985 29 51.8 108 2 G91297 regulator for trp
986 29 51.8 108 2 A86139 trp operon repress
987 29 51.8 109 2 S45257 trp operon repress
988 29 51.8 114 2 A95025 ribosomal protein
989 29 51.8 114 2 B97896 50S ribosomal prot
990 29 51.8 115 2 F86886 50S ribosomal prot
991 29 51.8 115 2 G95084 conserved domain p
992 29 51.8 115 2 C97952 hypothetical prote
993 29 51.8 117 2 T16313 hypothetical prote
994 29 51.8 121 2 G11919 pulI protein - Kle
995 29 51.8 125 1 A46315 E4 protein - human
996 29 51.8 126 2 E84047 hypothetical prote
997 29 51.8 132 1 S15618 E4 protein - human
998 29 51.8 137 2 A80751 probable lipoprote
999 29 51.8 137 2 E64956 yedD protein - Esc
1000 29 51.8 137 2 C85810 hypothetical prote

ALIGNMENTS

RESULT 1
T11560
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)
A:Species: simian immunodeficiency virus SIVsm
A:Variety: strain E543
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11560
R:Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.;
J. Virol. 71, 1609-1620, 1997
A:Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficie
A:Reference number: 217285; MUID:97151152; PMID:8995688
A:Accession: T11560
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1019 <HIR>
A:Cross-references: UNIPROT:P89154; UNIPARC:UPI0000105470; EMBL:U72748; NID:g1695908; P
C:Genetics: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; immunodeficiency

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 184 EGPKLRLQW 191
RESULT 2
AG0147
probable membrane protein YPO1203 [imported] - Yersinia pestis (strain CO92)
A:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 05-Oct-2004
C:Accession: AG0147
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0147
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <KUR>

A:Cross-references: UNIPROT:Q8ZGS7; UNIPARC:UPI00000DC87B; GB:AL590842; PIDN:CAC90042.1;
C:Genetics:
A:Gene: YPO1203
C:Superfamily: hypothetical protein yed
Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 PTLRQWLA 10
Db 66 PTLRQWAA 73
RESULT 3
S74539
hypothetical protein slr0740 - Synechocystis sp. (strain PCC 6803)
A:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74539
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KAN>
A:Cross-references: UNIPROT:P72684; UNIPARC:UPI00000C09C4; EMBL:D90899; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: Synechocystis hypothetical protein slr0740

Query Match 69.6%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
Db 29 PQLRQWLA 36

RESULT 4

S46354
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)
A:Species: simian immunodeficiency virus SIVagm
A:Variety: isolate SABD37
C:Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C:Accession: S46354
R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.; F
EMBO J. 13, 2935-2947, 1994
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A:Reference number: S46335; MUID:94298785; PMID:8026477
A:Accession: S46354
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-217 <JIN>

A:Cross-references: UNIPARC:UPI000010A53B; EMBL:U04018; NID:g466250; PIDN:AAA21512.1; P
A:Experimental source: isolate SABD37; saabaes monkey
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: polypotein

Query Match 69.6%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
: || || || ||

```
Db      86  DGPRLQW 93

RESULT 5
E84853
hypothetical protein At2g42400 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84853
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84853
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <STO>
A:Cross-references: UNIPROT:Q9SLB9; UNIPARC:UPI000017A02D; GB:AE002093; NID:G4567312; PI
C:Genetics:
A:Gene: At2g42400
A:Map position: 2

Query Match      69.6%; Score 39; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1  EGPTRLQWL 9
      :|||:|
Db      344  EGETIREWL 352

RESULT 6
S30484
pol polyprotein - human immunodeficiency virus type 2
C:Species: human immunodeficiency virus type 2, HIV-2
C>Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A:Reference number: S30460
A:Accession: S30484
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-656 <GAO>
A:Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
C:Superfamily: pol polyprotein

Query Match      69.6%; Score 39; DB 2; Length 656;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  EGPTRLQW 8
      :|||:|
Db      30  DGPRLQW 37

RESULT 7
S30483
pol polyprotein - human immunodeficiency virus type 2
C:Species: human immunodeficiency virus type 2, HIV-2
C>Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
C:Accession: S30483
R:Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
submitted to the EMBL Data Library, December 1992
A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
A:Reference number: S30460
A:Accession: S30483
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-656 <GAO>

A:Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111
C:Superfamily: pol polyprotein

Query Match      69.6%; Score 39; DB 2; Length 656;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  EGPTRLQW 8
      :|||:|
Db      30  DGPRLQW 37

RESULT 8
GNLJCA
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate CAM2/GUI)
N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly.
C:Species: human immunodeficiency virus type 2, HIV-2
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: B38475; JQ0974
R:Tristem, M.; Hill, P.; Karpas, A.
J. Gen. Virol. 72, 721-724, 1991
A:Title: Nucleotide sequence of a Guinea-Bissau-derived human immunodeficiency virus typ
A:Reference number: A38475; MUID:91170959; PMID:2005437
A:Accession: B38475
A:Molecule type: DNA
A:Residues: 1-1034 <TRI>
A:Cross-references: UNIPROT:P24107; UNIPARC:UPI0000131EF4
A:Note: readthrough of the terminator TGA may occur between codons ATT for 564-Ile and G
C:Comment: The cleavage sites of this polyprotein have not been determined.
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleo
F:85-183/Product: retropepsin #status predicted <RTP>
F:109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      69.6%; Score 39; DB 1; Length 1034;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  EGPTRLQW 8
      :|||:|
Db      200  DGPRLQW 207

RESULT 9
GNLJGG
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate GH-1)
N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C:Species: human immunodeficiency virus type 2, HIV-2
A:Note: host Homo sapiens (man)
C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004
C:Accession: JS0328
R:Hasegawa, A.; Tsujimoto, H.; Maki, N.; Ishikawa, K.; Miura, T.; Fukasawa, M.; Miki, K.
AIDS Res. Hum. Retroviruses 5, 593-604, 1989
A:Title: Sequence of a distinct HIV-2 isolate from Ghana showing significant divergence
A:Reference number: JS0327; MUID:90122350; PMID:2611042
A:Accession: JS0328
A:Molecule type: DNA
A:Residues: 1-1035 <HAS>
A:Cross-references: UNIPROT:P18042; UNIPARC:UPI0000174A34
A:Note: this sequence was submitted to JIPID, October 1989
C:Comment: Cleavage sites that yield the mature proteins remain to be determined.
C:Genetics:
A:Gene: pol
A:Start codon: ACA
C:Superfamily: pol polyprotein
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase;
F:85-183/Product: retropepsin #status predicted <RTP>
F:109/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      69.6%; Score 39; DB 1; Length 1035;
```

```
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
   :||| |||
Db 200 DGPRLRQW 207

RESULT 10
GNLJG2
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ROD)
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: B26262
R;Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M.
Nature 326, 662-669, 1987
A;Title: Genome organization and transactivation of the human immuno-deficiency virus ty
A;Reference number: A26262; MUID:87173056; PMID:3031510
A;Accession: B26262
A;Molecule type: DNA
A;Residues: 1-1036 <GUY>
A;Cross-references: UNIPROT:P04584; UNIPARC:UPI00000131EFA; GB:M15390
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C;Genetics:
A;Gene: pol
C;Superfamily: pol polypeptide
C;Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucle
F;110/Product: retropepsin #status predicted <RTP>
F;110/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 69.6%; Score 39; DB 1; Length 1036;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
   :||| |||
Db 201 DGPRLRQW 208

RESULT 11
GNLJST
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 2 (isolate ST)
N;Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C;Species: human immunodeficiency virus type 2, HIV-2
C;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: B33943
R;Kumar, P.; Hui, H.; Kappes, J.C.; Haggarty, B.S.; Hoxie, J.A.; Arya, S.K.; Shaw, G.M.;
J. Virol. 64, 890-901, 1990
A;Title: Molecular characterization of an attenuated human immunodeficiency virus type 2
A;Reference number: A33943; MUID:90112662; PMID:2296086
A;Accession: B33943
A;Molecule type: Genomic RNA
A;Residues: 1-1055 <KUM>
A;Cross-references: UNIPROT:P20876; UNIPARC:UPI00000131EFC
C;Comment: The pol polypeptide contains reverse transcriptase and endonuclease. However,
C;Genetics:
A;Gene: pol
C;Superfamily: pol polypeptide
C;Keywords: aspartic proteinase; endonuclease; hydrolase; nucleotidyltransferase; revers
F;105-203/Product: retropepsin #status predicted <RTP>
F;129/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 69.6%; Score 39; DB 1; Length 1055;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
   :||| |||
Db 220 DGPRLRQW 227
```

```
RESULT 12
S53092
pol polypeptide - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53092
R;Becker, M.; Zorr, B.; Becker, A.; Habermehl, K.O.
submitted to the EMBL Data Library, March 1995
A;Description: Molecular and phylogenetic characterisation of a Guinea-Bissau-derived hu
A;Reference number: S53091
A;Accession: S53092
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1055 <BEC>
A;Cross-references: UNIPROT:O73194; UNIPARC:UPI00001785DA; EMBL:Z48731
C;Superfamily: pol polypeptide
C;Keywords: polypeptide

Query Match 69.6%; Score 39; DB 2; Length 1055;
Best Local Similarity 75.0%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
   :||| |||
Db 220 DGPRLRQW 227

RESULT 13
B75468
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: B75468
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: B75468
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-134 <WHI>
A;Cross-references: UNIPROT:Q9RW17; UNIPARC:UPI00000C1826; GB:AE001939; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0852
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0852

Query Match 67.9%; Score 38; DB 2; Length 134;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GPTLRQW 8
   :||| :||
Db 15 GPTIKQW 21

RESULT 14
CGHU2B
collagen alpha 2(IV) chain precursor - human
N;Alternate names: procollagen alpha 2(IV) chain
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A32024; S00007; S02624; S00246; S17678; S16911; B32117; S16877; S00165; S39
R;Hostikka, S.L.; Tryggvason, K.
J. Biol. Chem. 263, 19488-19493, 1988
A;Title: The complete primary structure of the alpha2 chain of human type IV collagen an
A;Reference number: A32024; MUID:89066769; PMID:3198637
A;Accession: A32024
A;Molecule type: mRNA
A;Residues: 1-1712 <HOS1>
```

A;Cross-references: UNIPROT:P08572; UNIPARC:UPI0000126D42; EMBL:J04210; EMBL:X05610; GB:R:Hostikka, S.L.; Kurkinen, M.; Tryggvason, K.
 A;Title: The alpha-1(I)-beta-1 integrin recognition site of the basement membrane collagen I (alpha-1(I) chain) is located in the basement membrane collagen I (alpha-1(I) chain) region.
 A;Reference number: S00007; MUID:87219158; PMID:3582677
 A;Accession: S00007
 A;Molecule type: mRNA
 A;Residues: 1254-1398, 'V', 1400-1712 <HOS2>
 A;Cross-references: UNIPARC:UPI0000173BE6; EMBL:J04210; EMBL:X05610; GB:M20753; NID:g295
 A;Note: 1399-ile was also found
 R:Hostikka, S.L.; Tryggvason, K.
 FEBS Lett. 224, 297-305, 1987
 A;Title: Extensive structural differences between genes for the alpha(1) and alpha(2) chains of type I collagen
 A;Reference number: S02624; MUID:88083553; PMID:2826228
 A;Accession: S02624
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1347-1350; 1377-1383; 1426-1432; 1465-1471; 1529-1535; 1625-1630 <HOS3>
 A;Cross-references: UNIPARC:UPI0000142E66; UNIPARC:UPI0000173BE7; UNIPARC:UPI0000173BE8;
 A;Note: complete nucleotide sequence not shown
 R:Brazel, D.; Pollner, R.; Oberbaumer, I.; Kuehn, K.
 Eur. J. Biochem. 172, 35-42, 1988
 A;Title: Human basement membrane collagen (type IV): the amino acid sequence of the alpha-1(I) chain
 A;Reference number: S00246; MUID:88151998; PMID:3345760
 A;Accession: S00246
 A;Molecule type: mRNA
 A;Residues: 1-682, 'G', 684-1043 <BRA>
 A;Cross-references: UNIPARC:UPI0000173BEC; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PIDN:CAA29076.1; PIDN:CAA29076.1; PIDN:CAA29076.1
 R:Oberbaumer, I.
 submitted to the EMBL Data Library, June 1987
 A;Reference number: S17678
 A;Accession: S17678
 A;Molecule type: mRNA
 A;Residues: 1-470, 'P', 472-682, 'G', 684-1043 <OBE>
 A;Cross-references: UNIPARC:UPI000016A709; EMBL:X05562; NID:g30075; PIDN:CAA29076.1; PIDN:CAA29076.1; PIDN:CAA29076.1; PIDN:CAA29076.1
 R:Poeschl, E.; Pollner, R.; Kuehn, K.
 EMBL J. 7, 2687-2695, 1988
 A;Title: The genes for the alpha1(IV) and alpha2(IV) chains of human basement membrane collagen (type IV) are located on human chromosome 12
 A;Reference number: S02738; MUID:89030632; PMID:2846280
 A;Accession: S16911
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-33 <POE>
 A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:X12784; GB:M36963; NID:g30072; PIDN:CAA3
 R:Soiainen, R.; Huotari, M.; Hostikka, S.L.; Prockop, D.J.; Tryggvason, K.
 J. Biol. Chem. 263, 17217-17220, 1988
 A;Title: The structural genes for alpha1 and alpha2 chains of human type IV collagen are located on human chromosome 12
 A;Reference number: A92690; MUID:89034231; PMID:3182844
 A;Accession: B32117
 A;Molecule type: DNA
 A;Residues: 1-33 <SOI1>
 A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; EMBL:J05039; NID:g180759; PIDN:A
 R:Soiainen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Tryggvason, K.
 J. Biol. Chem. 264, 13565-13571, 1989
 A;Title: Structural organization of the gene for the alpha-1 chain of human type IV collagen
 A;Reference number: S16876; MUID:89340433; PMID:2701944
 A;Accession: S16877
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-33 <SOI2>
 A;Cross-references: UNIPARC:UPI000016A6F3; EMBL:J04217; NID:g180759; PIDN:AAA53097.1; PIDN:AAA53097.1; PIDN:AAA53097.1; PIDN:AAA53097.1
 A;Note: this sequence was submitted to the EMBL Data Library, October 1988
 R:Siebold, B.; Qian, R.Q.; Glanville, R.W.; Hofmann, H.; Deutzmann, R.; Kuehn, K.
 Eur. J. Biochem. 168, 569-575, 1987
 A;Title: Construction of a model for the aggregation and cross-linking region (7S domain) of type I collagen
 A;Reference number: S00165; MUID:88029476; PMID:3117548
 A;Accession: S00165
 A;Molecule type: protein
 A;Residues: 37-247 <SIE1>
 A;Cross-references: UNIPARC:UPI0000173BED
 A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 175-Gly

R;Eble, J.A.; Golbik, R.; Mann, K.; Kuehn, K.
 EMBO J. 12, 4795-4802, 1993
 A;Title: The alpha-1(I)-beta-1 integrin recognition site of the basement membrane collagen I (alpha-1(I) chain) is located in the basement membrane collagen I (alpha-1(I) chain) region.
 A;Reference number: S39614; MUID:94038963; PMID:8223488
 A;Accession: S39615
 A;Molecule type: protein
 A;Residues: 407-570 <EBL>
 A;Cross-references: UNIPARC:UPI0000173BEE
 R:MacKnight, R.S.; Benson, V.A.; Lovello, K.T.; van der Rest, M.; Fietzek, P.P.
 Biochemistry 22, 4940-4948, 1983
 A;Title: Isolation and characterization of pepsin-solubilized human basement membrane (type I) collagen
 A;Reference number: S16910; MUID:84053346; PMID:6416291
 A;Accession: S16912
 A;Molecule type: protein
 A;Residues: 490-492, 'X', 494-496, 675-677, 'G', 679-680, 'G', 682, 684-685, 'P' <MAC>
 A;Cross-references: UNIPARC:UPI0000173BEF; UNIPARC:UPI0000173BF0
 A;Experimental source: placenta
 R:Glanville, R.W.; Rauter, A.
 Hoppe-Seyler's Z. Physiol. Chem. 362, 943-951, 1981
 A;Title: Pepsin fragments of human placental basement-membrane collagens showing interrupted alpha-1(I) chains
 A;Reference number: S16908; MUID:82005835; PMID:6792033
 A;Accession: B58517
 A;Molecule type: protein
 A;Residues: 490-492, 'X', 494-501, 'P', 503-507, 952-957, 'X', 959-966, 'X', 968, 984-986, 'X', 988-989 <GLA>
 A;Cross-references: UNIPARC:UPI0000173BF1; UNIPARC:UPI0000173BF2; UNIPARC:UPI0000173BF3;
 R:Killen, P.D.; Francomano, C.A.; Yamada, Y.; Modi, W.S.; O'Brien, S.J.
 Hum. Genet. 77, 318-324, 1987
 A;Title: Partial structure of the human alpha-2(IV) collagen chain and chromosomal localization of the gene
 A;Reference number: S01450; MUID:88085168; PMID:3692475
 A;Accession: S01450
 A;Molecule type: mRNA
 A;Residues: 1040, 'L', 1042-1398, 'V', 1400-1418, 'M', 1420-1635, 'V', 1637-1712 <KIL>
 A;Cross-references: UNIPARC:UPI0000072E68; EMBL:M24766; NID:g537328; PIDN:AAA52043.1; PIDN:AAA52043.1; PIDN:AAA52043.1
 R:Siebold, B.; Deutzmann, R.; Kuehn, K.
 Eur. J. Biochem. 176, 617-624, 1988
 A;Title: The arrangement of intra- and intermolecular disulfide bonds in the carboxyterminal region of type IV collagen
 A;Reference number: S02550; MUID:89005112; PMID:2844531
 A;Accession: S02550
 A;Molecule type: protein
 A;Residues: 1480-1535, 1545-1614, 1617-1662, 'H', 1664-1700, 'G', 1705-1708, 1710-1712 <SIE2>
 A;Cross-references: UNIPARC:UPI0000173BF6; UNIPARC:UPI0000173BF7; UNIPARC:UPI0000173BF8;
 A;Note: the sequence from Fig. 7 is inconsistent with that shown in Fig. 11 in having 17
 R:Myers, J.C.; Howard, P.S.; Jelen, A.M.; Dion, A.S.; Macarak, E.J.
 J. Biol. Chem. 262, 9231-9238, 1987
 A;Title: Duplication of type IV collagen COOH-terminal repeats and species-specific expression of the gene
 A;Reference number: A27114; MUID:87250571; PMID:2439508
 A;Accession: B27114
 A;Molecule type: mRNA
 A;Residues: 1486-1574, 'I', 1576-1712 <MYE>
 A;Cross-references: UNIPARC:UPI0000173BFB; EMBL:J02760; NID:g180425; PIDN:AAA58422.1; PIDN:AAA58422.1; PIDN:AAA58422.1
 C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (Gly-X-Y) are substituted by O-glycosylated.
 C;Genetics:
 A;Gene: GDB:COL4A2
 A;Cross-references: GDB:119792; OMIM:120090
 A;Map position: 13q34-13q34
 A;Introns: 15/2, 33/3, 1347/1, 1380/1, 1429/1, 1468/1, 1532/1, 1532/3 #status incomplete
 A;Note: the alpha 1(IV) and alpha 2(IV) chain genes are encoded on opposite strands with
 C;Complex: Type IV collagen is a heterotrimer of two alpha 1(IV) chains (see PIR:CGH4B) and two alpha 2(IV) chains (see PIR:CGH4B) domains (with disulfide and desmosine cross-links), dimeric associations among trimer chains (with disulfide and desmosine cross-links), dimeric associations among trimer chains (with disulfide and desmosine cross-links).
 C;Function:
 A;Description: structural component of basement membrane
 C;Superfamily: collagen alpha 1(IV) chain
 C;Keywords: basement membrane; cell binding; coiled coil; extracellular matrix; glycoprotein; integrin; signal sequence #status predicted <SIG>
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-1712/Product: collagen alpha 2(IV) chain #status predicted <MAT>
 F:29-57/Domain: amino-terminal nonhelical, NH1 <NH1>
 F:58-1485/Region: interrupted helical
 F:362-364/Region: cell attachment (R-G-D) motif
 F:784-786/Region: cell attachment (R-G-D) motif
 F:868-870/Region: cell attachment (R-G-D) motif

F;889-891/Region: cell attachment (R-G-D) motif
F;970-972/Region: cell attachment (R-G-D) motif
F;1069-1071/Region: cell attachment (R-G-D) motif
F;1228-1230/Region: cell attachment (R-G-D) motif
F;1452-1454/Region: cell attachment (R-G-D) motif
F;1486-1712/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1495-1593/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1503-1708/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;442,47,51,53,137,483,485/Diulfide bonds: interchain #status predicted
F;57,87,90,102,165,168,225,239,242/Binding site: carboxydrate (Lys) #status atypical
F;57/Modified site: 5-hydroxylysine (Lys) #status atypical
F;63,75,96,114,123,132,150,159,186,189,198,201,213,216,219,496,499,955,964,1103,1115
F;87,90,102,165,168,225,239,242/Modified site: 5-hydroxylysine (Lys) #status experimental
F;138/Binding site: carboxydrate (Asn) (covalent) #status experimental
F;209/Modified site: 4-hydroxyproline (Pro) #status atypical
F;661-681/Diulfide bonds: #status predicted
F;1275/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;1504-1590,1537-1593/Diulfide bonds: (or 1504-1593, 1537-1590) #status experimental
F;1549-1555,1658-1665/Diulfide bonds: #status experimental
F;1612-1705,1646-1708/Diulfide bonds: (or 1612-1708, 1646-1705) #status experimental

Query Match 67.9%; Score 38; DB 1; Length 1712;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
||| |||
Db 10 GPDLRRWL 17

RESULT 15
A36925
transcription activator LysR-type Cbbr - Xanthobacter flavus
C;Species: Xanthobacter flavus
C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 09-Jul-2004
C;Accession: A36925; S13578; S35408
R;van den Bergh, B.R.E.; Dijkhuizen, L.; Meijer, W.G.
J. Bacteriol. 175, 6097-6104, 1993
A;Title: Cbbr, a LysR-type transcriptional activator, is required for expression of the
A;Reference number: A36925; MUID:94012468; PMID:8407781
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <VAN>
A;Cross-references: UNIPROT:P25545; UNIPARC:UPI0000127169; EMBL:Z22705; NID:G297851; PID
R;Meijer, W.G.; Arnborg, A.C.; Enquist, H.G.; Terpstra, P.; Lidstrom, M.E.; Dijkhuizen,
Mol. Gen. Genet. 225, 320-330, 1991
A;Title: Identification and organization of carbon dioxide fixation genes in Xanthobacte
A;Reference number: S13573; MUID:91172133; PMID:1900916
A;Accession: S13578
A;Molecule type: DNA
A;Residues: 1-150 <MEI>
A;Cross-references: UNIPARC:UPI00001788AC; EMBL:X17252
C;Genetics:
A;Gene: cbbr
A;Start codon: GTG
C;Superfamily: transcription activator LysR-type
C;Keywords: DNA binding; transcription regulation

Query Match 67.0%; Score 37.5; DB 2; Length 333;
Best Local Similarity 72.7%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 EG-PTLRQWLA 10
||| |||
Db 265 EGPLRVQWLA 275

RESULT 16
D70601
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galu [similarity] - Mycobacteri
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004

C;Accession: D70601
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70601
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-306 <COL>
A;Cross-references: UNIPROT:O05576; UNIPARC:UPI000000CCAD8; GB:Z94752; GB:AL123456; NID:G
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: galu
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidytransferase

Query Match 66.1%; Score 37; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
||| |||
Db 290 GPDLRRWL 297

RESULT 17
T45453
UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) galu [similarity] - Mycobacteri
C;Species: Mycobacterium leprae
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 05-Oct-2004
C;Accession: T45453
R;James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z22967
A;Accession: T45453
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-306 <JAM>
A;Cross-references: UNIPROT:Q9Z5G1; UNIPARC:UPI00000D438E; EMBL:AL035500; PIDN:CAB36696.1
A;Experimental source: cosmid L373
C;Genetics:
A;Note: galu
C;Superfamily: UDP-glucose pyrophosphorylase
C;Keywords: nucleotidytransferase

Query Match 66.1%; Score 37; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 9
||| |||
Db 290 GPDLRRWL 297

RESULT 18
T36550
hypothetical protein SCH10.38c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36550
R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21609
A;Accession: T36550
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-547 <MUR>
A;Cross-references: UNIPROT:Q9X8S7; UNIPARC:UPI00000DB04C; EMBL:AL049754; PIDN:CAB42047.1
A;Experimental source: strain A3(2)
C;Genetics:

A:Gene: SCOEDB:SCH10.38C

Query Match 66.1%; Score 37; DB 2; Length 547;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
|:|:|:|:
Db 57 PSLRLWLA 64

RESULT 19

A87309
hypothetical protein CC0482 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87309
R:NIERMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HEIDELBERG, J.
B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GWINN, M.L.; HAFT, D.H.; KOLON
N, J.; ERMOLAIEVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87309
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-777 <STO>
A:Cross-references: UNIPROT:Q9AAW1; UNIPARC:UPI00000C706B; GB:AE005673; NID:gl3421659; F
C:Genetics: CC0482
A:Gene: CC0482
C:Superfamily: cobalamin-independent methionine synthase

Query Match 66.1%; Score 37; DB 2; Length 777;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
|:|:|:|:
Db 361 PEIRQWLA 368

RESULT 20

A71006
hypothetical protein PH1343 - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: A71006
R:KAWARABAYASI, Y.; SAWADA, M.; HORIKAWA, H.; HAIKAWA, Y.; HINO, Y.; YAMAMOTO, S.; SEKII
M.; OHFUKU, Y.; FUNAHASHI, T.; TANAKA, T.; KUDOH, Y.; YAMAZAKI, J.; KUSHIDA, N.; OGUCHI
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71006
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-816 <KAW>
A:Cross-references: UNIPROT:O59068; UNIPARC:UPI0000063062; GB:AP0000006; NID:g3236133; PI
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1343

Query Match 66.1%; Score 37; DB 2; Length 816;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
|:|:|:|:
Db 114 QGOTLRNMLS 123

RESULT 21

JC4163

DNA-binding protein SE5 - rat
C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C:Accession: JC4163; PC4040
R:SUZUKI, E.; KOJIMA, N.; YOSHIMURA, K.; UYEMURA, K.; OBATA, K.; AKAGAWA, K.
J. Biochem. 118, 122-128, 1995
A:Title: Cloning and sequence analysis of cDNA for a possible DNA-binding protein SE5 in
A:Reference number: JC4163; MUID:96015159; PMID:8537300
A:Accession: JC4163
A:Molecule type: mRNA
A:Residues: 1-825 <SUZ>
A:Cross-references: UNIPROT:Q63003; UNIPARC:UPI0000124F19; DDBJ:D37934; NID:9531260; PID:
A:Experimental source: brain
A:Accession: PC4040
A:Molecule type: protein
A:Residues: 230-455 <SU2>
A:Cross-references: UNIPARC:UPI0000177392
C:Comment: This protein has an abundance of arginine, a glycine-rich region and a prolin
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: nerve; phosphoprotein
F:436-443/Region: nuclear location signal
F:722-731/Region: proline cluster
F:62,258,345,360,404,413,570,613,635,752,820/Binding site: phosphate (Ser) (covalent) (b

Query Match 66.1%; Score 37; DB 2; Length 825;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

Qy 1 EGPT--LRQWL 9
|:|:|:|:
Db 88 EGPSLPLRQWL 98

RESULT 22

S46347
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SAB-1)
C:Species: simian immunodeficiency virus SIVagm
A:Variety: isolate SAB-1
C:Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C:Accession: S46347
R:JIN, M.J.; HUI, H.; ROBERTSON, D.L.; MUELLER, M.C.; BARRE-SINOUSSE, F.; HIRSCH, V.M.;
EMBO J. 13, 2935-2947, 1994
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African gree
A:Reference number: S46335; MUID:94298785; PMID:8026477
A:Accession: S46347
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1039 <JIN>
A:Cross-references: UNIPARC:UPI00001096DD; EMBL:U04005; NID:9466229; PIDN:AAA21505.1; PI
A:Experimental source: isolate SAB-1; sabaesus monkey
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
A:Note: this reading frame extends between two stop codons and does not begin with a sta
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein

Query Match 66.1%; Score 37; DB 2; Length 1039;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
|:|:|:|:
Db 204 DGPRIRQW 211

RESULT 23

S08436
pol polyprotein - human immunodeficiency virus type 2 D205 (fragment)
C:Species: human immunodeficiency virus type 2 D205, HIV-2 D205
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S08436
R:DIETRICH, U.; ADAMSKI, M.; KREUTZ, R.; SEIPD, A.; KUEHNEL, H.; RUEBEAMEN-WAIGMANN, H.
Nature 342, 948-950, 1989

A;Title: A highly divergent HIV-2-related isolate.
A;Reference number: S08434; MUID:90081881; PMID:2594088
A;Accession: S08436
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1058 <DIE>
A;Cross-references: UNIPROT:P15833; UNIPARC:UPI0000131EF6; EMBL:X16109
A;Note: this sequence was submitted to the EMBL Data Library, Aug-1989
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: polyprotein

Query Match 66.1%; Score 37; DB 2; Length 1058;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8

Db 221 DGPKIRQW 228

RESULT 24

S54293
regulator protein p122-RhoGAP - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 01-Sep-1995
C;Accession: S54293
R;Homma, Y.; Emori, Y.
EMBO J. 14, 286-291, 1995

A;Title: A dual functional signal mediator showing RhoGAP and phospholipase C-delta stim

A;Reference number: S54293
A;Accession: S54293
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1083 <HOM>
A;Cross-references: UNIPARC:UPI000017C9AD; EMBL:D31962

Query Match 66.1%; Score 37; DB 2; Length 1083;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 916 EGPPRLRWRA 925

RESULT 25

S63748
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate 2-33-109
N;Contains: retropepsin (EC 3.4.23.16); RNA-directed DNA polymerase (EC 2.7.7.49)
C;Species: human immunodeficiency virus type 1, HIV-1
A;Variety: isolate 2-33-109
C;Date: 20-Jul-1996 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C;Accession: S63748; S63722

R;Yamaguchi, K.
submitted to the EMBL Data Library, July 1995

A;Reference number: S63731

A;Accession: S63748

A;Molecule type: DNA

A;Residues: 1-151 <YAM>

A;Cross-references: UNIPROT:Q72880; UNIPARC:UPI00000FFD38; EMBL:U31404; NID:g961586; PID

A;Experimental source: isolate 2-33-109

R;Yamaguchi, K.; Byrn, R.A.

Biochim. Biophys. Acta 1253, 136-140, 1995

A;Title: Clinical isolates of HIV-1 contain few pre-existing proteinase inhibitor resist

A;Reference number: S63703; MUID:96106422; PMID:8519793

A;Accession: S63722

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 23-121 <YAW>

A;Cross-references: UNIPARC:UPI0000108B23; EMBL:U31404

A;Experimental source: isolate 2-33-109

C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: AIDS; aspartic proteinase; hydrolase; immunodeficiency; nucleotidyltransferase

F;23-121/Product: retropepsin #status predicted <RTP>

Query Match 64.3%; Score 36; DB 2; Length 151;
Best Local Similarity 62.5%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8

Db 138 DGPKVRQW 145

RESULT 26

T23485

hypothetical protein K08F4.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C;Accession: T23485

R;Hembry, C.

submitted to the EMBL Data Library, January 1996

A;Reference number: Z19746

A;Accession: T23485

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-200 <WIL>

A;Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN00022; (

A;Experimental source: clone K08F4

C;Genetics:
A;Gene: CRSP:K08F4.11

A;Map position: 4

A;Introns: 45/1; 76/1; 111/3

C;Superfamily: glutathione transferase

Query Match 64.3%; Score 36; DB 2; Length 200;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 184 ETPKLEWLA 193

RESULT 27

T37464

probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T37464

R;Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.

submitted to the EMBL Data Library, June 1997

A;Description: Paraquat mediates differential gene expression in C. elegans.

A;Reference number: Z21702

A;Accession: T37464

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-207 <TAW>

A;Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AA65419.1

A;Experimental source: strain Bristol N2

C;Genetics:
A;Gene: GST3

C;Superfamily: glutathione transferase

C;Keywords: transferase

Query Match 64.3%; Score 36; DB 2; Length 207;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10

Db 191 ETPKLEWLA 200


```
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
Db 376 GPTLDKWL 383

RESULT 33
A:Reference: A81958
A:Title: probable permease NMA0414 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
A:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: A81958
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jogle, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: A81958
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-530 <PAR>
A:Cross-references: UNIPROT:Q9UWE3; UNIPARC:UPI000000C497F; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0414

Query Match 64.3%; Score 36; DB 2; Length 530;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db 191 EMPVLRPWLA 200

RESULT 34
A:Reference: B81015
A:Title: ABC transporter, permease protein NMB2026 [imported] - Neisseria meningitidis (strain MC
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81015
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81015
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-531 <TGT>
A:Cross-references: UNIPROT:Q9JXI9; UNIPARC:UPI000000C4849; GB:AE002552; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB2026

Query Match 64.3%; Score 36; DB 2; Length 531;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EGPTLRQWLA 10
Db 192 EMPVLRPWLA 201

RESULT 35
A:Reference: S29955
A:Title: surface glycoprotein - feline immunodeficiency virus
A:Species: feline immunodeficiency virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S29955
```

```
R;Rigby, M.A.
submitted to the EMBL Data Library, November 1992
A:Reference number: S29946
A:Accession: S29955
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-827 <RIG>
A:Cross-references: UNIPARC:UPI0000170E56; EMBL:X59496; NID:g397348; PIDN:CAA49250.1; PFI
C:Superfamily: feline immunodeficiency virus env polyprotein
C:Keywords: glycoprotein

Query Match 64.3%; Score 36; DB 2; Length 827;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 197 EGPTLGW 204

RESULT 36
A:Reference: JQ2003
A:Title: env polyprotein - feline immunodeficiency virus (strain UK2)
N:Alternate names: coat polyprotein
N:Contains: surface glycoprotein; transmembrane glycoprotein
C:Species: feline immunodeficiency virus
A:Note: host Felis silvestris catus (domestic cat)
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2003; S29954
R;Rigby, M.A.; Holmes, E.C.; Pistello, M.; Mackay, A.; Brown, A.J.L.; Neil, J.C.
J. Gen. Virol. 74, 425-436, 1993
A:Title: Evolution of structural proteins of feline immunodeficiency virus: Molecular ep
A:Reference number: JQ2003; MUID:93187604; PMID:8383177
A:Accession: JQ2003
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-855 <RIG>
A:Cross-references: UNIPROT:Q04993; UNIPARC:UPI0000129FFD; GB:X69494; NID:g59267; PIDN:CH
C:Genetics:
A:Gene: env
C:Superfamily: feline immunodeficiency virus env polyprotein
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei
F:1-610/Product: surface glycoprotein #status predicted <SPR>
F:95-111/Region: hydrophobic
F:151-178/Region: hydrophobic
F:611-855/Product: transmembrane glycoprotein #status predicted <TMP>
F:611-648/Region: hydrophobic
F:786-802/Domain: transmembrane #status predicted <TMN>
F:258,269,274,298,418,548,551,716,720,728,736/Binding site: carbohydrate (Asn) (covalent)

Query Match 64.3%; Score 36; DB 1; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 225 EGPTLGW 232

RESULT 37
A:Reference: JQ2004
A:Title: env polyprotein - feline immunodeficiency virus (strain UK8)
N:Alternate names: coat polyprotein
N:Contains: surface glycoprotein; transmembrane glycoprotein
C:Species: feline immunodeficiency virus
A:Note: host Felis silvestris catus (domestic cat)
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2004
R;Rigby, M.A.; Holmes, E.C.; Pistello, M.; Mackay, A.; Brown, A.J.L.; Neil, J.C.
J. Gen. Virol. 74, 425-436, 1993
A:Title: Evolution of structural proteins of feline immunodeficiency virus: Molecular ep
A:Reference number: JQ2003; MUID:93187604; PMID:8383177
A:Accession: JQ2004
```

A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-855 <RG>
A;Cross-references: UNIPROT:Q04995; UNIPARC:UPI0000129FFE; GB:X69496
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-610/Product: surface glycoprotein #status predicted <SPR>
F;95-111/Region: hydrophobic
F;151-172/Region: hydrophobic
F;611-855/Product: transmembrane glycoprotein #status predicted <TMP>
F;611-648/Region: hydrophobic
F;786-802/Domain: transmembrane #status predicted <TMN>
F;258,269,274,298,418,548,551,716,720,728,736/Binding site: carbohydrate (Asn) (covalent)

Query Match 64.3%; Score 36; DB 1; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 38
F45557
external glycoprotein, transmembrane protein(external glycoprotein, EGP, transmembrane protein)
C;Species: feline immunodeficiency virus
C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: F45557
R;Maki, N.; Miyazawa, T.; Fukasawa, M.; Hasegawa, A.; Hayami, M.; Miki, K.; Mikami, T.
Arch. Virol. 123, 29-45, 1992
A;Title: Molecular characterization and heterogeneity of feline immunodeficiency virus isolates
A;Reference number: A45557; MUID:92198230; PMID:1312825
A;Accession: F45557
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-855 <MA>
A;Cross-references: UNIPROT:Q02282; UNIPARC:UPI0000129FFC
A;Experimental source: strain TM2
A;Note: sequence extracted from NCBI backbone (NCBI:89826, NCBIP:89854)
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: glycoprotein; transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 855;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 39
A45394
env polyprotein precursor - feline immunodeficiency virus (strain UT-113)
N;Alternate names: coat polyprotein
N;Contains: surface glycoprotein; transmembrane glycoprotein
C;Species: feline immunodeficiency virus
A;Note: host Felis silvestris catus (domestic cat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A45394; S16030
R;Verschoor, E.J.; Hulakotte, E.G.J.; Ederveen, J.; Koolen, M.J.M.; Horzinek, M.C.; Rott
Virology 193, 433-438, 1993
A;Title: Post-translational processing of the feline immunodeficiency virus envelope protein
A;Reference number: A45394; MUID:93174954; PMID:8382405
A;Accession: A45394
A;Molecule type: mRNA
A;Residues: 1-856 <VR>
A;Cross-references: UNIPROT:Q03804; UNIPARC:UPI000011D647; EMBL:X60725; NID:g1092; PIDN:
C;Comment: This protein lacks an N-terminal signal sequence, and one of the three internal
C;Genetics:

A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-611/Product: surface glycoprotein #status predicted <SGP>
F;95-111/Region: hydrophobic
F;151-169/Region: hydrophobic
F;612-856/Product: transmembrane glycoprotein #status predicted <TGP>
F;616-640/Region: hydrophobic
F;786-812/Domain: transmembrane #status predicted <TMN>
F;220,258,269,274,298,330,336,342,418,422,448,469,481,499,518,531,548,551,556,717,721,722,737,742

Query Match 64.3%; Score 36; DB 1; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 40
VCLJFP
env polyprotein precursor - feline immunodeficiency virus (strain Petaluma)
N;Alternate names: coat polyprotein
N;Contains: surface glycoprotein; transmembrane glycoprotein
C;Species: feline immunodeficiency virus
A;Note: host Felis silvestris catus (domestic cat)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: P33543
R;Talbot, R.L.; Sparger, E.E.; Lovelace, K.M.; Fitch, W.M.; Pedersen, N.C.; Luciw, P.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 5743-5747, 1989
A;Title: Nucleotide sequence and genomic organization of feline immunodeficiency virus.
A;Reference number: A33543; MUID:89345543; PMID:2762293
A;Accession: P33543
A;Molecule type: DNA
A;Residues: 1-856 <TAL>
A;Cross-references: UNIPROT:P16090; UNIPARC:UPI0000129PFA; GB:M25381
C;Comment: This protein lacks an N-terminal signal sequence, and one of the three internal
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-611/Product: surface glycoprotein #status predicted <SGP>
F;95-111/Region: hydrophobic
F;151-169/Region: hydrophobic
F;612-856/Product: transmembrane glycoprotein #status predicted <TGP>
F;616-640/Region: hydrophobic
F;786-812/Domain: transmembrane #status predicted <TMN>
F;220,258,269,274,298,330,336,342,418,422,448,481,499,518,531,548,551,717,721,729,737/B1

Query Match 64.3%; Score 36; DB 1; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
| | | | |
Db 225 EGPTLGNW 232

RESULT 41
S23823
env polyprotein - feline immunodeficiency virus
C;Species: feline immunodeficiency virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S23823
R;Morikawa, S.; Lutz, H.; Bishop, D.H.L.
submitted to the EMBL Data Library, December 1990
A;Description: Identification of conserved and variable regions in the envelope glycoprotein
A;Reference number: S19886
A;Accession: S23823
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-856 <MOR>

```
A;Cross-references: UNIPROT:Q66975; UNIPARC:UPI000010A513; EMBL:X57002; NID:G59285; PIDN
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: polyprotein

Query Match      64.3%; Score 36; DB 2; Length 856;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 225 EGPTLGNN 232

RESULT 42
S19886
env polyprotein E - feline immunodeficiency virus
N/Alternate names: envelope polyprotein E
C;Species: feline immunodeficiency virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S19886
R;Morikawa, S.; Lutz, H.; Bishop, D.H.L.
A;Submitted to the EMBL Data Library, December 1990
A;Description: Identification of conserved and variable regions in the envelope glycoprotein
A;Reference number: S19886
A;Accession: S19886
A;Molecule type: genomic RNA
A;Status: preliminary
A;Residues: 1-857 <MOR>
A;Cross-references: UNIPROT:Q66954; UNIPARC:UPI000010E436; EMBL:X57001; NID:G59279; PIDN
C;Genetics:
A;Gene: env
C;Superfamily: feline immunodeficiency virus env polyprotein
C;Keywords: glycoprotein; polyprotein

Query Match      64.3%; Score 36; DB 2; Length 857;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 224 EGPTLGNN 231

RESULT 43
AB1885
hypothetical protein alr0627 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB1885
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saemoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB1885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <KUR>
A;Cross-references: UNIPROT:Q8VZ60; UNIPARC:UPI00000CDDDB; GB:BA0000019; PIDN:BA072585.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0627

Query Match      64.3%; Score 36; DB 2; Length 907;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
Db 107 GPTLMSWL 114

RESULT 44
```

```
S12153
pol polyprotein - human immunodeficiency virus type 2
C;Species: human immunodeficiency virus type 2, HIV-2
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S12153
R;Kuehnelt, H.; Kreutz, R.; Ruebsamen-Waigmann, H.
Nucleic Acids Res. 18, 6142, 1990
A;Title: Nucleotide sequence of HIV-2(D194), an isolate from a Gambian case of 'Neuro-AII
A;Reference number: S12152; MUID:91045094; PMID:2235509
A;Accession: S12153
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1032 <KUE>
A;Cross-references: UNIPROT:P17757; UNIPARC:UPI00001785D6; EMBL:X52223
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1990
C;Superfamily: pol polyprotein

Query Match      64.3%; Score 36; DB 2; Length 1032;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 197 DGPRLXQW 204

RESULT 45
GNLJG3
HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (macaque isolate)
N/Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym
C;Species: simian immunodeficiency virus, SIV
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
C;Accession: B28887
R;Chakrabarti, L.; Guyader, M.; Alizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.;
Nature 328, 543-547, 1987
A;Title: Sequence of simian immunodeficiency virus from macaque and its relationship to
A;Reference number: A28887; MUID:87287230; PMID:3649576
A;Accession: B28887
A;Molecule type: DNA
A;Residues: 1-1056 <CHA>
A;Cross-references: UNIPARC:UPI0000131F1B; GB:Y00277; GB:M16403; NID:G61730; PIDN:CAA683
C;Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev
C;Genetics:
A;Gene: pol
C;Superfamily: pol polyprotein
C;Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse
P;106-204/Product: retropepsin #status predicted <RTP>
P;130/Active site: Asp (shared with dimeric partner) #status predicted

Query Match      64.3%; Score 36; DB 1; Length 1056;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 221 DGPRLXQW 228

RESULT 46
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N/Alternate names: protein F5E19.190
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51517
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1123 <SAT>
A;Cross-references: UNIPROT:Q9SPU7; UNIPARC:UPI00000A945F; EMBL:AL391147
```

A:Experimental source: cultivar Columbia; BAC clone F5E19

C:Genetics:

A:Map position: 5

A:Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2

A:Note: F5E19_190

Query Match 64.3%; Score 36; DB 2; Length 1123;

Best Local Similarity 75.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10

||| |||

Db 203 PTKRQWLS 210

RESULT 47

S28764

A:Experimental source: rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C:Accession: S28764

R:Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.

J. Biol. Chem. 267, 19536-19547, 1992

A:Title: Cloning and primary structure of neurocan, a developmentally regulated, aggregated

A:Reference number: S28764; MUID:92406907; PMID:1326557

A:Accession: S28764

A:Molecule type: mRNA

A:Residues: 1-1257 <RAU>

A:Cross-references: UNIPROT:P55067; UNIPARC:UPI0000131788; EMBL:M97161; NID:G205649; PID

C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG

C:Keywords: chondroitin sulfate proteoglycan; glycoprotein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-1257/Product: neurocan #status predicted <MAT>

F:176-253/Domain: link protein repeat homology <LNK1>

F:274-355/Domain: link protein repeat homology <LNK2>

F:364-366/Region: cell attachment (R-G-D) motif

F:953-984/Domain: EGF homology <EGF>

F:1029-1149/Domain: C-type lectin homology <LCH>

F:1156-1212/Domain: complement factor H repeat homology <FHD>

F:1121,339,737,967,1164/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:372,410/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

F:944/Binding site: chondroitin sulfate (Ser) (covalent) #status experimental

Query Match 64.3%; Score 36; DB 2; Length 1257;

Best Local Similarity 75.0%; Pred. No. 2.8e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9

||| |||

Db 434 GPTLASWL 441

RESULT 48

T40245

A:Experimental source: Schizosaccharomyces pombe

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T40245

R:Bozzym, K.; Beck, A.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1999

A:Reference number: Z21916

A:Accession: T40245

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1339 <BOR>

A:Cross-references: UNIPROT:Q9UST7; UNIPARC:UPI0000069A03; EMBL:AL121815; PIDN:CAB58159.

A:Experimental source: strain 972h-; cosmid c336

C:Genetics:

A:Gene: SPDB:SPBC336.07

A:Map position: 2

Query Match 64.3%; Score 36; DB 2; Length 1339;

Best Local Similarity 85.7%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQW 8

||| |||

Db 334 GPTLPQW 340

RESULT 49

T45278

A:oligopeptide ABC transport protein bidKB [imported] - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004

C:Accession: T45278

R:Nodwell J., R.; McGovern, K.; Losick, R.

submitted to the EMBL Data Library, August 1996

A:Description: An oligopeptide permease responsible for the import of an extracellular

A:Reference number: Z22954

A:Accession: T45278

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-602 <NOD>

A:Cross-references: UNIPROT:P72407; UNIPARC:UPI00000BC8D6; EMBL:U68036; PIDN:AAB09555.1

A:Experimental source: strain M145

C:Genetics:

A:Note: bldKB

C:Function:

A:Description: involved in aerial mycelium formation

C:Keywords: oligopeptide transport

Query Match 63.4%; Score 35.5; DB 2; Length 602;

Best Local Similarity 63.6%; Pred. No. 1.6e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Qy 1 EGPT-LRQWLA 10

||| |||

Db 174 DGPTYLQWLS 184

RESULT 50

A97655

A:hypochemical protein (U34346) [imported] - Agrobacterium tumefaciens (strain C58, Cerson

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C:Accession: A97655

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, B.;

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: A97655

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-120 <KUR>

A:Cross-references: UNIPROT:Q8UCM6; UNIPARC:UPI00000D1BEC; GB:AE007869; PIDN:AAK88194.1.

C:Genetics:

A:Gene: AGR_C 4462

A:Map position: circular chromosome

Query Match 62.5%; Score 35; DB 2; Length 120;

Best Local Similarity 71.4%; Pred. No. 37;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GPTLRQW 8

||| |||

Db 58 GPTSLRRW 64

Search completed: May 12, 2006, 10:52:17

Job time : 17.2368 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 70.7692 seconds
(without alignments)

99.694 Million cell updates/sec

Title: US-10-632-388-62

Perfect score: 56

Sequence: 1 EGPTLRQWLA 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	78.6	319	Q9RKM5_STRCO	Q9RKM5 streptomyce
2	43	76.8	297	Q7UQ84_RHOBA	Q7UQ84 rhodospirill
3	42	75.0	313	P90433_SIVCZ	P90433 chimpanzee
4	42	75.0	580	Q89RH2_BRAJA	Q89RH2 bradyrhizob
5	42	75.0	1019	POL_SIVS4	P12502 simian immu
6	42	75.0	1019	P89154_SIVCZ	P89154 chimpanzee
7	42	75.0	1019	Q72BR7_SIVCZ	Q72BR7 chimpanzee
8	42	75.0	1019	Q7ZBR5_SIVCZ	Q7ZBR5 chimpanzee
9	41	73.2	244	Q9R7K1_9SPHN	Q9R7K1 erythroba
10	41	73.2	245	Q66272_9SPHN	Q66272 erythroba
11	41	73.2	245	Q82991_9SPHN	Q82991 erythroba
12	41	73.2	249	Q82989_9SPHN	Q82989 erythroba
13	41	73.2	278	Q9XDV0_9SPHN	Q9XDV0 erythroba
14	41	73.2	302	Q742B3_MYCPA	Q742B3 mycobacteri
15	41	73.2	308	XERC_CORGL	Q8nm29 corynebacte
16	41	73.2	728	Q617Q3_CAEER	Q617Q3 caenorhabdi
17	41	73.2	754	Q95Y82_CAEER	Q95Y82 caenorhabdi
18	41	73.2	791	Q73SJ6_MYCPA	Q73SJ6 mycobacteri
19	41	73.2	815	Q4KS46_VIRU	Q4KS46 orange-spot
20	41	73.2	863	Q8ST50_MAIZE	Q8ST50 zea mays (m
21	41	73.2	941	Q8QUJ6_VIRU	Q8QUJ6 infectious
22	40	71.4	53	Q5YJ55_NOCFA	Q5YJ55 nocardia fa
23	40	71.4	129	Q8DHX7_SYNEL	Q8DHX7 synechococc
24	40	71.4	137	Q6FE11_ACIAD	Q6FE11 acinetobact
25	40	71.4	154	Q54XF9_DICDI	Q54XF9 dictyosteli
26	40	71.4	296	Q8ZGS7_YERPE	Q8ZGS7 yersinia pe
27	40	71.4	296	Q66D06_YERPS	Q66D06 yersinia ps
28	40	71.4	326	P95613_RHIGA	P95613 rhizobium g
29	40	71.4	377	Q82PX5_STRAW	Q82PX5 streptomyce
30	40	71.4	429	Q7S2Z7_NEUCR	Q7S2Z7 neurospora
31	40	71.4	760	Q7G4N4_ORISA	Q7G4N4 oryza sativ

32	39	69.6	131	2	P72684_SYNY3	P72684 synechocyst
33	39	69.6	133	2	Q4ZPV7_PSEBSY	Q4ZPV7 pseudomonas
34	39	69.6	134	2	Q86G15_PSEBSM	Q86G15 pseudomonas
35	39	69.6	154	2	Q9WI01_9HIV2	Q9WI01 human immu
36	39	69.6	164	2	Q6VNC9_BDRBA	Q6VNC9 bdellovibri
37	39	69.6	168	2	Q9V492_DROME	Q9V492 drosophila
38	39	69.6	217	2	Q87115_SIVCZ	Q87115 chimpanzee
39	39	69.6	218	2	Q5RJN8_RAT	Q5RJN8 rattus norv
40	39	69.6	230	2	Q7ZKL1_9HIV1	Q7ZKL1 human immu
41	39	69.6	245	2	Q82987_9SPHN	Q82987 erythroba
42	39	69.6	298	2	Q5YPT6_NOCFA	Q5YPT6 nocardia fa
43	39	69.6	310	2	Q9D8Z7_MOUSE	Q9D8Z7 mus musculu
44	39	69.6	310	2	Q7SK19_9HIV2	Q7SK19 human immu
45	39	69.6	338	2	Q4PIZ3_USTWA	Q4PIZ3 utellago ma
46	39	69.6	340	2	Q8UN04_SIVCZ	Q8UN04 chimpanzee
47	39	69.6	340	2	Q8UN03_SIVCZ	Q8UN03 chimpanzee
48	39	69.6	349	2	Q7SKK9_9HIV2	Q7SKK9 human immu
49	39	69.6	349	2	Q7SKK8_9HIV2	Q7SKK8 human immu
50	39	69.6	349	2	Q7SKK6_9HIV2	Q7SKK6 human immu
51	39	69.6	349	2	Q7SKK4_9HIV2	Q7SKK4 human immu
52	39	69.6	349	2	Q7SKJ6_9HIV2	Q7SKJ6 human immu
53	39	69.6	349	2	Q7SKJ5_9HIV2	Q7SKJ5 human immu
54	39	69.6	349	2	Q7SKJ4_9HIV2	Q7SKJ4 human immu
55	39	69.6	349	2	Q7SKJ3_9HIV2	Q7SKJ3 human immu
56	39	69.6	349	2	Q7SKJ2_9HIV2	Q7SKJ2 human immu
57	39	69.6	349	2	Q7SKJ1_9HIV2	Q7SKJ1 human immu
58	39	69.6	349	2	Q7SKJ0_9HIV2	Q7SKJ0 human immu
59	39	69.6	349	2	Q7SKI8_9HIV2	Q7SKI8 human immu
60	39	69.6	349	2	Q7SKI7_9HIV2	Q7SKI7 human immu
61	39	69.6	349	2	Q7SKI6_9HIV2	Q7SKI6 human immu
62	39	69.6	349	2	Q7SKI5_9HIV2	Q7SKI5 human immu
63	39	69.6	349	2	Q7SKI4_9HIV2	Q7SKI4 human immu
64	39	69.6	349	2	Q7SKI3_9HIV2	Q7SKI3 human immu
65	39	69.6	349	2	Q7SKI2_9HIV2	Q7SKI2 human immu
66	39	69.6	349	2	Q7SKI1_9HIV2	Q7SKI1 human immu
67	39	69.6	351	2	Q9E3T1_9HIV2	Q9E3T1 human immu
68	39	69.6	351	2	Q9E3T2_9HIV2	Q9E3T2 human immu
69	39	69.6	351	2	Q9E3T4_9HIV2	Q9E3T4 human immu
70	39	69.6	396	2	Q90PUI_SIVCZ	Q90PUI chimpanzee
71	39	69.6	450	2	Q9SLB9_ARATH	Q9SLB9 arabidopsis
72	39	69.6	472	2	Q9WHZ9_9HIV2	Q9WHZ9 human immu
73	39	69.6	476	2	Q9WI00_9HIV2	Q9WI00 human immu
74	39	69.6	496	2	Q6CER7_YARLI	Q6CER7 yarrowia li
75	39	69.6	560	2	Q06347_9HIV2	Q06347 human immu
76	39	69.6	586	2	Q828W4_STRAW	Q828W4 streptomyce
77	39	69.6	593	2	C5TNY2_ANOGA	C5TNY2 anopheles g
78	39	69.6	740	2	Q8XVE2_RALSO	Q8XVE2 ralstonia s
79	39	69.6	756	2	Q4ZQB5_PSEBSY	Q4ZQB5 pseudomonas
80	39	69.6	756	2	Q885P2_PSEBSM	Q885P2 pseudomonas
81	39	69.6	986	2	C57059_SIVCZ	C57059 chimpanzee
82	39	69.6	1017	2	Q6VG40_SIVCZ	Q6VG40 chimpanzee
83	39	69.6	1022	1	POL_SIVSP	P19505 simian immu
84	39	69.6	1022	2	Q90317_SIVCZ	Q90317 chimpanzee
85	39	69.6	1022	2	Q87956_SIVCZ	Q87956 chimpanzee
86	39	69.6	1022	2	Q87965_SIVCZ	Q87965 chimpanzee
87	39	69.6	1022	2	Q89620_SIVCZ	Q89620 chimpanzee
88	39	69.6	1022	2	Q88135_SIVCZ	Q88135 chimpanzee
89	39	69.6	1034	1	POL_HV2CA	P24107 human immu
90	39	69.6	1035	2	Q7J194_9HIV2	Q7J194 human immu
91	39	69.6	1035	2	Q6R793_9HIV2	Q6R793 human immu
92	39	69.6	1035	2	Q6R784_9HIV2	Q6R784 human immu
93	39	69.6	1036	1	POL_HV2RO	P18042 human immu
94	39	69.6	1049	1	POL_HV2G1	P18042 human immu
95	39	69.6	1055	1	POL_HV2ST	P20876 human immu
96	39	69.6	1055	2	Q9YTU0_9HIV2	Q9YTU0 human immu
97	39	69.6	1056	2	Q04097_SIVCZ	Q04097 chimpanzee
98	39	69.6	1059	2	Q76630_9HIV2	Q76630 human immu
99	39	69.6	1060	2	Q5QG01_SIVCZ	Q5QG01 chimpanzee
100	39	69.6	1060	2	Q5QFT5_SIVCZ	Q5QFT5 chimpanzee
101	39	69.6	1060	2	Q5QFR3_SIVCZ	Q5QFR3 chimpanzee
102	39	69.6	1060	2	Q5QFL2_SIVCZ	Q5QFL2 chimpanzee
103	39	69.6	1060	2	Q5QFJ6_SIVCZ	Q5QFJ6 chimpanzee
104	39	69.6	1209	2	Q4WNF8_ASPFU	Q4WNF8 aspergillus

105	39	69.6	1226	2	Q6H0K6_9HIV2	Q6H0K6 human immun	178	37	66.1	998	2	Q8JBD2_9HIV1	Q8JBD2 human immun
106	38	69.6	1414	2	Q7S131_NEUCR	Q7S131 neurospora	179	37	66.1	1015	2	Q8JAH1_SIVCZ	Q8JAH1 chimpanzee
107	38	67.9	134	2	Q9RW17_DEIRA	Q9RW17 deirococcus	180	37	66.1	1016	2	Q8JAI0_SIVCZ	Q8JAI0 chimpanzee
108	38	67.9	177	2	Q4UTU1_XANCP	Q4UTU1 xanthomonas	181	37	66.1	1018	2	Q9IBL9_9HIV2	Q9IBL9 human immun
109	38	67.9	177	2	Q8PA24_XANCP	Q8PA24 xanthomonas	182	37	66.1	1039	2	Q87I08_SIVCZ	Q87I08 chimpanzee
110	38	67.9	244	2	Q66269_9SPHN	Q66269 erythronici	183	37	66.1	1055	2	Q89928_9HIV2	Q89928 human immun
111	38	67.9	245	2	Q66278_9SPHN	Q66278 agrobacteri	184	37	66.1	1058	1	POL HV2D2	P15833 human immun
112	38	67.9	245	2	Q9ZN87_9SPHN	Q9ZN87 porphyrobac	185	37	66.1	1061	2	Q76634_9HIV2	Q76634 human immun
113	38	67.9	246	2	Q66276_9SPHN	Q66276 porphyrobac	186	37	66.1	1064	2	Q8UTV7_9HIV2	Q8UTV7 human immun
114	38	67.9	316	2	Q5FPU4_GLUOX	Q5FPU4 gluconobact	187	37	66.1	1091	1	RHG07_FAT	Q63744 r rho-glcpas
115	38	67.9	375	2	Q7FANA_ORYSA	Q7FANA oryza sativ	188	37	66.1	1157	2	Q527F0_MAGGR	Q527F0 magnaporthe
116	38	67.9	375	2	Q7XPP6_ORYSA	Q7XPP6 oryza sativ	189	37	66.1	1357	2	Q5AXV0_EMENI	Q5AXV0 aspergillus
117	38	67.9	403	2	Q5DN43_9CAUD	Q5DN43 bacterioph	190	37	66.1	1391	2	Q4N6K3_THEPA	Q4N6K3 theileria p
118	38	67.9	403	2	Q88NU2_PSEPK	Q88NU2 pseudomonas	191	37	66.1	1413	2	Q4UJ10_THEAN	Q4UJ10 theileria a
119	38	67.9	476	2	Q4WL94_ASPTU	Q4WL94 aspergillus	192	37	66.1	2355	2	Q4KCD0_PSEPF	Q4KCD0 pseudomonas
120	38	67.9	631	2	Q6LP74_PHOPR	Q6LP74 photobacter	193	36	64.3	85	2	Q64GN7_9HIV1	Q64GN7 human immun
121	38	67.9	818	2	Q7WYU0_ALCEU	Q7WYU0 alcaligenes	194	36	64.3	114	2	Q730Z9_BACC1	Q730Z9 bacillus ce
122	38	67.9	1235	2	Q52FK0_MAGGR	Q52FK0 magnaporthe	195	36	64.3	115	2	Q81M37_BACAN	Q81M37 bacillus an
123	38	67.9	1304	2	Q7SDK1_NEUCR	Q7SDK1 neurospora	196	36	64.3	127	2	Q97396_9HIV1	Q97396 human immun
124	38	67.9	1319	2	Q61EX3_CAEBR	Q61EX3 caenorhabdi	197	36	64.3	151	2	Q72880_9HIV1	Q72880 human immun
125	38	67.9	1391	2	Q5AVS1_EMENI	Q5AVS1 aspergillus	198	36	64.3	152	2	Q8Q9F0_9HIV1	Q8Q9F0 human immun
126	38	67.9	1706	2	Q5VZA9_HUMAN	Q5VZA9 homo sapien	199	36	64.3	152	2	Q8Q9F3_9HIV1	Q8Q9F3 human immun
127	38	67.9	1712	1	Q4A2A2_HUMAN	Q4A2A2 homo sapien	200	36	64.3	166	2	Q6TIB6_9HIV1	Q6TIB6 human immun
128	37.5	67.0	333	1	Q5V0V6_CAEEL	Q5V0V6 caenorhabdi	202	36	64.3	176	2	Q656A5_ORYSA	Q656A5 oryza sativ
129	37	66.1	103	2	Q950V6_CAEEL	Q950V6 caenorhabdi	203	36	64.3	203	2	Q4VTU5_9SAUR	Q4VTU5 gymmodactyl
130	37	66.1	129	2	Q4I3M8_GIBZE	Q4I3M8 gibberella	204	36	64.3	207	1	GST3 CAEEL	Q16116 caenorhabdi
131	37	66.1	132	2	Q8NI04_HUMAN	Q8NI04 homo sapien	205	36	64.3	213	2	Q4VTU4_9SAUR	Q4VTU4 gymmodactyl
132	37	66.1	153	2	Q80ZRO_MOUSE	Q80ZRO mus musculu	206	36	64.3	221	2	Q8M2E3_9SAUR	Q8M2E3 pachydyactyl
133	37	66.1	168	2	Q9B596_CHILA	Q9B596 chinchilla	207	36	64.3	221	2	Q8M2D8_9SAUR	Q8M2D8 pachydyactyl
134	37	66.1	182	2	Q9B598_CHILA	Q9B598 chinchilla	208	36	64.3	226	2	Q8Q3M7_9HIV1	Q8Q3M7 human immun
135	37	66.1	192	2	Q7NT62_CHRVO	Q7NT62 chromobacte	209	36	64.3	229	2	Q9WKB0_9HIV1	Q9WKB0 human immun
136	37	66.1	192	2	Q95519_NEORB	Q95519 indirana sp	209	36	64.3	233	2	Q9QDG2_9HIV1	Q9QDG2 human immun
137	37	66.1	209	2	Q6N1X5_RHOPA	Q6N1X5 rhodopseudo	210	36	64.3	238	2	Q4VTU3_9SAUR	Q4VTU3 gymmodactyl
138	37	66.1	238	2	Q835U7_ENTFA	Q835U7 enterococcu	211	36	64.3	239	2	Q5MI13_9HIV1	Q5MI13 human immun
139	37	66.1	238	2	Q12158_9HIV1	Q12158 human immun	212	36	64.3	240	2	Q5MI07_9HIV1	Q5MI07 human immun
140	37	66.1	270	2	Q4SAB1_TETNG	Q4SAB1 tetraodon n	213	36	64.3	240	2	Q509K7_9HIV1	Q509K7 human immun
141	37	66.1	297	2	Q03307_CHELO	Q03307 chelodina l	214	36	64.3	242	2	Q7PIP3_CHRVO	Q7PIP3 chromobacte
142	37	66.1	297	2	Q03335_9SAUR	Q03335 pelusios wi	215	36	64.3	243	2	Q75801_9HIV1	Q75801 human immun
143	37	66.1	300	2	Q5RPD1_9HIV1	Q5RPD1 human immun	216	36	64.3	244	2	Q4V8S1_BRARE	Q4V8S1 brachydanio
144	37	66.1	301	2	Q4NHT3_9MICC	Q4NHT3 arthrobacte	217	36	64.3	245	2	Q6QTJ9_9HIV1	Q6QTJ9 human immun
145	37	66.1	306	2	Q7D906_MYCTU	Q7D906 mycobacteri	218	36	64.3	247	2	Q75802_9HIV1	Q75802 human immun
146	37	66.1	306	2	Q05576_MYCTU	Q05576 mycobacteri	219	36	64.3	248	2	Q6EPI6_9HIV1	Q6EPI6 human immun
147	37	66.1	306	2	Q7U0W3_MYCBO	Q7U0W3 mycobacteri	220	36	64.3	250	2	Q903C2_9HIV1	Q903C2 human immun
148	37	66.1	306	2	Q9Z5G1_MYCLE	Q9Z5G1 mycobacteri	221	36	64.3	253	2	Q6LXI1_PHOPR	Q6LXI1 photobacter
149	37	66.1	326	2	Q98VC4_SHEPC	Q98VC4 hepatitis c	222	36	64.3	255	2	Q5MXX9_9HIV1	Q5MXX9 human immun
150	37	66.1	329	2	Q9DZC1_9HIV1	Q9DZC1 human immun	223	36	64.3	255	2	Q7ZKF3_9HIV1	Q7ZKF3 human immun
151	37	66.1	329	2	Q9DZC2_9HIV1	Q9DZC2 human immun	224	36	64.3	261	2	Q5RRV0_9HIV1	Q5RRV0 human immun
152	37	66.1	332	2	Q7ZMA0_9HIV1	Q7ZMA0 human immun	225	36	64.3	266	2	Q6EPL4_9HIV1	Q6EPL4 human immun
153	37	66.1	338	2	Q4LPN5_9HURK	Q4LPN5 burkholderi	226	36	64.3	267	2	Q7N2A0_PHOIL	Q7N2A0 photorhabdu
154	37	66.1	355	1	TSN10_HUMAN	Q8H129 homo sapien	227	36	64.3	278	2	Q56276_THIPE	Q56276 thibacillu
155	37	66.1	364	1	MRAN_BORBR	Q7Wf55 bordetella	228	36	64.3	278	2	Q9KWI4_ERYLO	Q9KWI4 erythroba
156	37	66.1	364	1	MRAN_BORPA	Q7W4a7 bordetella	229	36	64.3	280	2	Q32396_ERYLO	Q32396 erythroba
157	37	66.1	369	1	MRAN_BORPE	Q7Vup6 bordetella	230	36	64.3	284	2	Q9RJ55_STRCO	Q9RJ55 streptomyce
158	37	66.1	369	2	Q69LD7_ORYSA	Q69LD7 oryza sativ	231	36	64.3	284	2	Q8ENM1_OCBTH	Q8ENM1 oceanobacil
159	37	66.1	382	2	Q8U337_PVRPU	Q8U337 pyrococcus	232	36	64.3	285	2	Q6K4U4_ORYSA	Q6K4U4 oryza sativ
160	37	66.1	386	1	ETR1_CANTR	Q8Wzm3 candida tro	233	36	64.3	295	2	Q9ERG8_9HIV1	Q9ERG8 human immun
161	37	66.1	386	1	ETR2_CANTR	Q8Wzm4 candida tro	234	36	64.3	298	2	Q8J800_9HIV1	Q8J800 human immun
162	37	66.1	405	2	Q5LQ49_SILPO	Q5LQ49 silicibacte	235	36	64.3	299	2	Q5RMX3_9HIV1	Q5RMX3 human immun
163	37	66.1	414	2	Q4LO61_9HIV1	Q4LO61 human immun	236	36	64.3	300	2	Q6R664_9HIV1	Q6R664 human immun
164	37	66.1	423	2	Q8AH22_9HIV1	Q8AH22 human immun	237	36	64.3	300	2	Q5RRL4_9HIV1	Q5RRL4 human immun
165	37	66.1	481	2	Q8L117_RHOSH	Q8L117 rhodobacter	238	36	64.3	300	2	Q5RNL4_9HIV1	Q5RNL4 human immun
166	37	66.1	485	2	Q8DSX5_VIBVU	Q8DSX5 vibrio vuln	239	36	64.3	301	2	Q8L0P1_MOUSE	Q8L0P1 mus musculu
167	37	66.1	547	2	Q9X8S7_STRCO	Q9X8S7 streptomyce	240	36	64.3	302	2	Q8KUC3_RHRET	Q8KUC3 rhizobium e
168	37	66.1	632	2	Q75I09_ORYSA	Q75I09 oryza sativ	241	36	64.3	303	2	Q8J821_9HIV1	Q8J821 human immun
169	37	66.1	669	2	Q526R7_MAGGR	Q526R7 magnaporthe	242	36	64.3	304	2	Q5RER7_9HIV1	Q5RER7 human immun
170	37	66.1	733	2	Q5XQF4_MAIZE	Q5XQF4 sea mays (m	243	36	64.3	307	2	Q9LT41_ARATH	Q9LT41 arabidopsis
171	37	66.1	750	2	Q7UUG1_RHOBA	Q7UUG1 rhodopirell	244	36	64.3	308	2	Q6R6F9_9HIV1	Q6R6F9 human immun
172	37	66.1	777	1	METE_CAUCR	Q9aa1 caulobacter	245	36	64.3	312	2	Q9CE34_LACLA	Q9CE34 lactococcus
173	37	66.1	816	2	Q59088_FRRHO	Q59088 pyrococcus	246	36	64.3	315	2	Q8LF62_ARATH	Q8LF62 arabidopsis
174	37	66.1	818	2	Q5NV45_9BURK	Q5NV45 raiatonia m	247	36	64.3	316	2	Q8LWV5_ARATH	Q8LWV5 arabidopsis
175	37	66.1	825	1	SE5_RAT	Q6J499 rat	248	36	64.3	316	2	Q50A19_9HIV1	Q50A19 human immun
176	37	66.1	853	2	Q8Q5Q4_9RETR	Q8Q5Q4 feline immu	249	36	64.3	319	2		
177	37	66.1	854	2	Q6J4Y9_9RETR	Q6J4Y9 feline immu	250	36	64.3				

251	36	64.3	320	2	Q8YH8_9H1V1	Q8YH8 human immun	324	36	64.3	391	2	Q75P0_PHPA	Q75P0 physcomitre
252	36	64.3	321	2	Q8JB46_9H1V1	Q8JB46 human immun	325	36	64.3	391	2	Q8Y28_9H1V1	Q8Y28 human immun
253	36	64.3	321	2	Q8Q60_9H1V1	Q8Q60 human immun	326	36	64.3	392	2	Q3575_9H1V1	Q3575 human immun
254	36	64.3	324	1	Q51D1_HUMAN	Q8Q63 homo sapien	327	36	64.3	401	2	Q90IW3_9H1V1	Q90IW3 human immun
255	36	64.3	324	2	Q6WP7_9H1V1	Q6WP7 human immun	328	36	64.3	402	2	Q4TMT6_9SPHN	Q4TMT6 erythrobaet
256	36	64.3	326	2	Q41273_GIBBERELLA	Q41273 gibberella	329	36	64.3	402	2	Q82R87_STRAW	Q82R87 streptomyce
257	36	64.3	327	2	Q5BJV7_RAT	Q5BJV7 rattus norv	330	36	64.3	417	2	Q90MT8_9H1V1	Q90MT8 human immun
258	36	64.3	327	2	Q9D959_MOUSE	Q9D959 mus musculu	331	36	64.3	419	2	Q8AEC2_9H1V1	Q8AEC2 human immun
259	36	64.3	328	2	Q9ED10_9H1V1	Q9ED10 human immun	332	36	64.3	420	2	Q9P523_NEUCR	Q9P523 neurospora
260	36	64.3	328	2	Q9ECQ1_9H1V1	Q9ECQ1 human immun	333	36	64.3	421	2	Q7WQUB_BORDER	Q7WQUB bordetella
261	36	64.3	328	2	Q9EINI_9H1V1	Q9EINI human immun	334	36	64.3	421	2	Q9QCY1_9H1V1	Q9QCY1 human immun
262	36	64.3	328	2	Q9EE76_9H1V1	Q9EE76 human immun	335	36	64.3	423	2	Q8AHP1_9H1V1	Q8AHP1 human immun
263	36	64.3	328	2	Q9EEW3_9H1V1	Q9EEW3 human immun	336	36	64.3	423	2	Q8AH66_9H1V1	Q8AH66 human immun
264	36	64.3	328	2	Q9ECX4_9H1V1	Q9ECX4 human immun	337	36	64.3	423	2	Q8AH61_9H1V1	Q8AH61 human immun
265	36	64.3	328	2	Q9EC5_9H1V1	Q9EC5 human immun	338	36	64.3	425	2	Q4VZ57_9H1V1	Q4VZ57 human immun
266	36	64.3	328	2	Q9EKU1_9H1V1	Q9EKU1 human immun	339	36	64.3	426	2	Q9PXC7_9RETR	Q9PXC7 feline immu
267	36	64.3	328	2	Q9ECJ5_9H1V1	Q9ECJ5 human immun	340	36	64.3	432	2	Q4JK01_9H1V1	Q4JK01 human immun
268	36	64.3	328	2	Q9ELJ2_9H1V1	Q9ELJ2 human immun	341	36	64.3	434	2	Q5D836_9H1V1	Q5D836 human immun
269	36	64.3	328	2	Q9EDJ2_9H1V1	Q9EDJ2 human immun	342	36	64.3	434	2	Q5D836_9H1V1	Q5D836 human immun
270	36	64.3	328	2	Q9EH06_9H1V1	Q9EH06 human immun	343	36	64.3	457	1	Y4BF_RHISN	Y4BF rhodopirell
271	36	64.3	328	2	Q9ECF3_9H1V1	Q9ECF3 human immun	344	36	64.3	457	2	Q7ULE0_RHOBA	Q7ULE0 rhodopirell
272	36	64.3	328	2	Q9EGW3_9H1V1	Q9EGW3 human immun	345	36	64.3	458	2	Q5QW6_CRYNE	Q5QW6 cryptococcu
273	36	64.3	329	2	Q904D4_9H1V1	Q904D4 human immun	346	36	64.3	467	2	Q5Z206_NOCFA	Q5Z206 nocardia fa
274	36	64.3	329	2	Q8QBR7_9H1V1	Q8QBR7 human immun	347	36	64.3	468	2	Q88H13_PSEPK	Q88H13 pseudomonas
275	36	64.3	330	2	Q50A84_9H1V1	Q50A84 human immun	348	36	64.3	479	2	Q5D9T4_SCHJA	Q5D9T4 schistosoma
276	36	64.3	332	2	Q8Q9H6_9H1V1	Q8Q9H6 human immun	349	36	64.3	481	2	Q7RUZ6_NEUCR	Q7RUZ6 neurospora
277	36	64.3	332	2	Q8Q9J8_9H1V1	Q8Q9J8 human immun	350	36	64.3	499	2	Q67D18_9H1V1	Q67D18 human immun
278	36	64.3	332	2	Q8Q9H4_9H1V1	Q8Q9H4 human immun	351	36	64.3	499	2	Q67D97_9H1V1	Q67D97 human immun
279	36	64.3	332	2	Q8Q9H3_9H1V1	Q8Q9H3 human immun	352	36	64.3	499	2	Q67D81_9H1V1	Q67D81 human immun
280	36	64.3	332	2	Q8Q919_9H1V1	Q8Q919 human immun	353	36	64.3	499	2	Q67CP5_9H1V1	Q67CP5 human immun
281	36	64.3	332	2	Q8Q912_9H1V1	Q8Q912 human immun	354	36	64.3	499	2	Q67CB1_9H1V1	Q67CB1 human immun
282	36	64.3	332	2	Q7ZMT9_9H1V1	Q7ZMT9 human immun	355	36	64.3	499	2	Q67CA2_9H1V1	Q67CA2 human immun
283	36	64.3	332	2	Q7ZM65_9H1V1	Q7ZM65 human immun	356	36	64.3	499	2	Q511W4_9H1V1	Q511W4 human immun
284	36	64.3	332	2	Q7ZM00_9H1V1	Q7ZM00 human immun	357	36	64.3	499	2	Q511T3_9H1V1	Q511T3 human immun
285	36	64.3	332	2	Q7ZL28_9H1V1	Q7ZL28 human immun	358	36	64.3	504	2	Q6EMC3_ECOLI	Q6EMC3 escherichia
286	36	64.3	332	2	Q7ZLV1_9H1V1	Q7ZLV1 human immun	359	36	64.3	504	2	Q6WTH8_9ZZZZ	Q6WTH8 plasmid r10
287	36	64.3	333	2	Q8JAL6_9H1V1	Q8JAL6 human immun	360	36	64.3	507	1	TRA6_PSEAE	TRA6 pseudomonas
288	36	64.3	335	2	Q82D06_STRAW	Q82D06 streptomyce	361	36	64.3	507	1	TRA6_PSEAE	TRA6 pseudomonas
289	36	64.3	335	2	Q8QAP6_9H1V1	Q8QAP6 human immun	362	36	64.3	507	1	TRA6_PSEAE	TRA6 pseudomonas
290	36	64.3	335	2	Q8QAP8_9H1V1	Q8QAP8 human immun	363	36	64.3	507	2	Q7BT57_ECOLI	Q7BT57 escherichia
291	36	64.3	336	2	Q533H4_9H1V1	Q533H4 human immun	364	36	64.3	507	2	Q79FC0_PSEPN	Q79FC0 klebsiella
292	36	64.3	336	1	ASB1_MOUSE	Q9VW74 mus musculu	365	36	64.3	507	2	Q79FC0_PSEPN	Q79FC0 klebsiella
293	36	64.3	340	2	Q50B8_9H1V1	Q50B8 human immun	366	36	64.3	524	2	Q91DL5_9H1V1	Q91DL5 human immun
294	36	64.3	341	2	Q8MPQ3_BDEBA	Q8MPQ3 bdellovibri	367	36	64.3	524	2	Q91DL5_9H1V1	Q91DL5 human immun
295	36	64.3	342	2	Q80T19_MOUSE	Q80T19 mus musculu	368	36	64.3	530	2	Q5F58_NEISERIA	Q5F58 neisseria g
296	36	64.3	346	2	Q89M02_BRAJA	Q89M02 bradyrhizob	369	36	64.3	530	2	Q5F58_NEISERIA	Q5F58 neisseria m
297	36	64.3	349	2	Q90M47_9H1V1	Q90M47 human immun	370	36	64.3	531	2	Q9JWE3_NEIMA	Q9JWE3 neisseria m
298	36	64.3	349	2	Q7SKL0_9H1V2	Q7SKL0 human immun	371	36	64.3	531	2	Q9JX19_NEIMA	Q9JX19 neisseria m
299	36	64.3	349	2	Q7SKK7_9H1V2	Q7SKK7 human immun	372	36	64.3	532	2	Q5GW75_XANOR	Q5GW75 xanthomonas
300	36	64.3	349	2	Q7SKK9_9H1V2	Q7SKK9 human immun	373	36	64.3	542	2	Q4SIY0_TETNG	Q4SIY0 tetraodon n
301	36	64.3	349	2	Q7SKK3_9H1V2	Q7SKK3 human immun	374	36	64.3	542	2	Q4SIY0_TETNG	Q4SIY0 tetraodon n
302	36	64.3	349	2	Q7SKK2_9H1V2	Q7SKK2 human immun	375	36	64.3	559	2	Q99B10_9H1V1	Q99B10 human immun
303	36	64.3	349	2	Q7SKK1_9H1V2	Q7SKK1 human immun	376	36	64.3	562	2	Q99B10_9H1V1	Q99B10 human immun
304	36	64.3	349	2	Q7SKK0_9H1V2	Q7SKK0 human immun	377	36	64.3	562	2	Q4SJK6_TETNG	Q4SJK6 tetraodon n
305	36	64.3	349	2	Q7SKJ8_9H1V2	Q7SKJ8 human immun	378	36	64.3	566	2	Q4SJK6_TETNG	Q4SJK6 tetraodon n
306	36	64.3	349	2	Q7SKJ8_9H1V2	Q7SKJ8 human immun	379	36	64.3	571	2	Q7S508_NEUCR	Q7S508 neurospora
307	36	64.3	349	2	Q7SKJ8_9H1V2	Q7SKJ8 human immun	380	36	64.3	579	2	Q6WLD7_9BACT	Q6WLD7 uncultured
308	36	64.3	351	2	Q9E3T6_9H1V2	Q9E3T6 human immun	381	36	64.3	586	2	Q70ER5_9H1V1	Q70ER5 human immun
309	36	64.3	351	2	Q9E3T6_9H1V2	Q9E3T6 human immun	382	36	64.3	586	2	Q70ER5_9H1V1	Q70ER5 human immun
310	36	64.3	351	2	Q8Q9K2_9H1V1	Q8Q9K2 human immun	383	36	64.3	598	2	Q6N4Y3_RHOPA	Q6N4Y3 rhodopseudo
311	36	64.3	351	2	Q9E3T5_9H1V2	Q9E3T5 human immun	384	36	64.3	607	2	Q6N4Y3_RHOPA	Q6N4Y3 rhodopseudo
312	36	64.3	351	2	Q9E3T5_9H1V2	Q9E3T5 human immun	385	36	64.3	610	2	Q91L84_POLCB	Q91L84 polyangium
313	36	64.3	354	2	Q8AEQ9_STRCO	Q8AEQ9 human immun	386	36	64.3	627	2	Q4RLG7_TETNG	Q4RLG7 tetraodon n
314	36	64.3	356	2	Q8AEQ6_9H1V1	Q8AEQ6 human immun	387	36	64.3	630	2	Q4RLG7_TETNG	Q4RLG7 tetraodon n
315	36	64.3	360	2	Q90KA3_9H1V1	Q90KA3 human immun	388	36	64.3	633	2	Q61V52_9RETR	Q61V52 feline immu
316	36	64.3	367	2	Q5KFI9_CRYNE	Q5KFI9 cryptococcu	389	36	64.3	637	2	Q4SJK6_TETNG	Q4SJK6 tetraodon n
317	36	64.3	370	2	Q8UJB9_AGR75	Q8UJB9 agrobacteri	390	36	64.3	657	2	Q6JSB7_9H1V1	Q6JSB7 human immun
318	36	64.3	371	2	Q9K567_MYCAU	Q9K567 mycobacteri	391	36	64.3	657	2	Q6JSB7_9H1V1	Q6JSB7 human immun
319	36	64.3	375	2	Q7VUT6_BORPA	Q7VUT6 bordetella	392	36	64.3	657	2	Q6JSB7_9H1V1	Q6JSB7 human immun
320	36	64.3	375	2	Q7W3W9_BORPA	Q7W3W9 bordetella	393	36	64.3	657	2	Q6JSB7_9H1V1	Q6JSB7 human immun
321	36	64.3	375	2	Q7WF99_BORBR	Q7WF99 bordetella	394	36	64.3	661	2	Q6JSB7_9H1V1	Q6JSB7 human immun
322	36	64.3	384	2	Q9J3Z1_9H1V1	Q9J3Z1 human immun	395	36	64.3	674	2	Q9NDH7_CAEEL	Q9NDH7 caenorhabdi
323	36	64.3	387	2	Q9GCP0_9H1V1	Q9GCP0 human immun	396	36	64.3	699	2	Q9NDH7_CAEEL	Q9NDH7 caenorhabdi

397	36	64.3	764	2	Q4WK0 ASPFU	Q4wkm0 aspergillus	470	36	64.3	1060	2	Q5QFS2_S1VCZ	Q5qfs2 chimpanzee
398	36	64.3	799	2	Q82KF5_STRAW	Q82kfs streptomyce	471	36	64.3	1060	2	Q5QF04_S1VCZ	Q5qf04 chimpanzee
399	36	64.3	817	2	Q98KX1_RHILO	Q98kx1 rhizobium l	472	36	64.3	1060	2	Q5QFP4_S1VCZ	Q5qfp4 chimpanzee
400	36	64.3	830	2	Q521F6_MAGGR	Q521f6 magnaporthe	473	36	64.3	1060	2	Q5QFN4_S1VCZ	Q5qfn4 chimpanzee
401	36	64.3	851	2	Q8UNJ5_feline immu	Q8unj5 feline immu	474	36	64.3	1060	2	Q5QFM3_S1VCZ	Q5qfm3 chimpanzee
402	36	64.3	851	2	Q8UNJ6_9RETR	Q8unj6 feline immu	475	36	64.3	1060	2	Q5QFK8_S1VCZ	Q5qfk8 chimpanzee
403	36	64.3	852	2	Q66959_9RETR	Q66959 feline immu	476	36	64.3	1060	2	Q5IZC2_S1VCZ	Q5izc2 chimpanzee
404	36	64.3	853	2	Q66935_9RETR	Q66935 feline immu	477	36	64.3	1072	2	Q5BIM0_DROME	Q5bim0 drosophila
405	36	64.3	853	2	Q66957_9RETR	Q66957 feline immu	478	36	64.3	1072	2	Q9W592_DROME	Q9w592 drosophila
406	36	64.3	853	2	Q61V53_9RETR	Q61v53 feline immu	479	36	64.3	1073	1	POL_HV2D1	P17757 human immun
407	36	64.3	854	1	ENV_FIVSD	Q19030 feline immu	480	36	64.3	1090	2	Q4XZK5_PLACH	Q4xzks plasmodium
408	36	64.3	854	1	ENV_FIVMO	Q05312 feline immu	481	36	64.3	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis
409	36	64.3	854	2	Q90DK4_9RETR	Q90dk4 feline immu	482	36	64.3	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis
410	36	64.3	855	1	ENV_FIVt2	Q02282 feline immu	483	36	64.3	1142	1	POL_HV2BE	P18096 human immun
411	36	64.3	855	1	ENV_FIVU8	Q04993 feline immu	484	36	64.3	1257	1	CSFG3_RAT	P55067 rattus norv
412	36	64.3	855	1	ENV_FIVU8	Q04993 feline immu	485	36	64.3	1330	2	Q41254_GIBZE	Q41254 gibberella
413	36	64.3	855	2	Q66958_9RETR	Q66958 feline immu	486	36	64.3	1334	2	Q5XHG9_XENLA	Q5xhg9 xenopus lae
414	36	64.3	855	2	Q66960_9RETR	Q66960 feline immu	487	36	64.3	1335	1	RPTOR_HUMAN	Rhn122 homo sapien
415	36	64.3	855	2	Q66961_9RETR	Q66961 feline immu	488	36	64.3	1335	1	RPTOR_MOUSE	Q8k4g0 mus musculu
416	36	64.3	855	2	Q66962_9RETR	Q66962 feline immu	489	36	64.3	1339	2	Q9UST7_SCHPO	Q9ust7 schizosacch
417	36	64.3	855	2	Q61V18_9RETR	Q61v18 feline immu	490	36	64.3	1365	2	Q4SD20_TETNG	Q4sd20 tetraodon n
418	36	64.3	855	2	Q89617_9RETR	Q89617 feline immu	491	36	64.3	1431	2	Q6BQ71_DESHA	Q6bq71 debaryomyce
419	36	64.3	856	1	ENV_FIVPE	P16090 feline immu	492	36	64.3	1431	2	Q5B491_EMENI	Q5b491 aspergillus
420	36	64.3	856	1	ENV_FIVU1	Q03804 feline immu	493	36	64.3	1435	2	Q4WIA7_STRAW	Q4wia7 streptomyce
421	36	64.3	856	2	Q03800_9RETR	Q03800 feline immu	494	36	64.3	1514	2	Q82BH0_STRAW	Q82bh0 streptomyce
422	36	64.3	856	2	Q03801_9RETR	Q03801 feline immu	495	36	64.3	3165	2	Q5AR63_EMENI	Q5ar63 aspergillus
423	36	64.3	856	2	Q66975_9RETR	Q66975 feline immu	496	35.5	63.4	399	2	Q7VT67_BORPE	Q7vt67 bordetella
424	36	64.3	856	2	Q9PX25_9RETR	Q9px25 feline immu	497	35.5	63.4	399	2	Q7W289_BORPA	Q7w289 bordetella
425	36	64.3	857	2	Q66954_9RETR	Q66954 feline immu	498	35.5	63.4	399	2	Q7WR56_BORBR	Q7wr56 pseudomonas
426	36	64.3	857	2	Q66956_9RETR	Q66956 feline immu	499	35.5	63.4	531	2	Q87WE7_PSSSM	Q87we7 pseudomonas
427	36	64.3	859	2	Q90OK5_9RETR	Q90ok5 feline immu	500	35.5	63.4	600	2	Q93IU2_STRCO	Q93iu2 streptomyce
428	36	64.3	869	2	Q9VZ82_DROME	Q9vz82 drosophila	501	35.5	63.4	602	2	P72407_STRCO	P72407 streptomyce
429	36	64.3	876	2	P88142_9HIV2	P88142 human immun	502	35.5	63.4	705	2	Q6WEF8_PECCE	Q6wef8 peccobacter
430	36	64.3	898	2	Q4RLG8_TETNG	Q4rlg8 tetraodon n	503	35	62.5	65	2	Q984B5_RHILO	Q984b5 rhizobium l
431	36	64.3	907	2	Q8YZ60_ANASP	Q8yz60 anabaena sp	504	35	62.5	68	2	Q7R2G9_GIALA	Q7r2g9 giardia lam
432	36	64.3	912	2	Q9Q622_9HIV1	Q9q622 human immun	505	35	62.5	72	2	Q6K895_ORYSA	Q6k895 oryza sativ
433	36	64.3	936	2	Q7ZJ29_S1VCZ	Q7zj29 simian immu	506	35	62.5	85	2	Q64GR5_9HIV1	Q64gr5 human immun
434	36	64.3	954	2	Q651T2_ORYSA	Q651t2 oryza sativ	507	35	62.5	87	2	Q65930_CAEV	Q65930 caprine art
435	36	64.3	999	2	Q8UTG9_9HIV1	Q8utg9 human immun	508	35	62.5	87	2	Q9WR47_9RETR	Q9wr47 brasilian c
436	36	64.3	1002	2	Q4QX95_9HIV1	Q4qx95 human immun	509	35	62.5	91	2	Q8Y015_RALSO	Q8y015 ralstonia s
437	36	64.3	1003	2	Q6PR23_9HIV1	Q6pr23 human immun	510	35	62.5	110	2	Q856U7_9CAUD	Q856u7 mycobacteri
438	36	64.3	1003	2	Q519G4_9HIV1	Q519g4 human immun	511	35	62.5	117	2	Q4TNK7_9SPHN	Q4tnk7 erythrobact
439	36	64.3	1005	2	Q6Y8X5_9HIV1	Q6y8x5 human immun	512	35	62.5	120	2	Q8UCM6_AGRT5	Q8ucm6 agrobacteri
440	36	64.3	1006	2	Q5U8G6_9HIV1	Q5u8g6 human immun	513	35	62.5	124	2	Q5FUP7_GLOJX	Q5fup7 gluconobact
441	36	64.3	1006	2	Q596M2_9HIV1	Q596m2 human immun	514	35	62.5	144	2	Q7U677_SYNPX	Q7u677 synecococc
442	36	64.3	1007	2	Q66QF6_9HIV1	Q66qf6 human immun	515	35	62.5	148	2	Q4UK00_RICPE	Q4uk00 rickettsia
443	36	64.3	1035	1	POL_HV2KR	Q74120 human immun	516	35	62.5	160	2	Q73UZ7_MYCPA	Q73uz7 mycobacteri
444	36	64.3	1035	1	POL_HV2NZ	P05952 human immun	517	35	62.5	173	2	Q8PLV3_XANAC	Q8plv3 xanthomonas
445	36	64.3	1043	2	Q90DD2_S1VCZ	Q90dd2 chimpanzee	518	35	62.5	173	2	Q9B593_9HYST	Q9b593 lagidium vi
446	36	64.3	1056	1	POL_SIVM1	P05896 simian immu	519	35	62.5	182	2	Q94Q15_CHIBR	Q94q15 chinchilla
447	36	64.3	1057	1	POL_SIVAI	Q02836 simian immu	520	35	62.5	182	2	Q9B594_9HYST	Q9b594 lagidium vi
448	36	64.3	1059	2	Q56317_9PLVG	Q56317 simian-huma	521	35	62.5	190	2	Q9BHP3_ARATH	Q9bhp3 arabidopsis
449	36	64.3	1059	2	O11404_9PLVG	O11404 simian-huma	522	35	62.5	195	2	Q6N108_RHOPA	Q6n108 rhodospseudo
450	36	64.3	1059	2	O56320_9PLVG	O56320 simian-huma	523	35	62.5	197	1	BETI_PSEAE	Q9ht10 pseudomonas
451	36	64.3	1059	2	O56613_9PLVG	O56613 simian-huma	524	35	62.5	197	1	BETI_PSESM	Q8aif0 pseudomonas
452	36	64.3	1060	2	Q90EX3_S1VCZ	Q90ex3 simian immu	525	35	62.5	197	2	Q4ZM61_PSSSY	Q4zm61 pseudomonas
453	36	64.3	1060	2	Q88016_S1VCZ	Q88016 chimpanzee	526	35	62.5	197	2	Q4K4K9_PSPF5	Q4k4k9 pseudomonas
454	36	64.3	1060	2	Q90EX8_S1VCZ	Q90ex8 simian immu	527	35	62.5	197	2	Q92NS0_RHIME	Q92ns0 rhizobium m
455	36	64.3	1060	2	Q8USP8_9PLVG	Q8usp8 simian-huma	528	35	62.5	206	2	Q8PKQ0_XANAC	Q8pkq0 xanthomonas
456	36	64.3	1060	2	Q07387_S1VCZ	Q07387 chimpanzee	529	35	62.5	226	2	Q7NFQ5_GLOVI	Q7nfq5 gloebacter
457	36	64.3	1060	2	Q5W9H4_S1VCZ	Q5w9h4 chimpanzee	530	35	62.5	232	2	Q509W0_9HIV1	Q509w0 human immun
458	36	64.3	1060	2	Q5QFP0_S1VCZ	Q5qfp0 chimpanzee	531	35	62.5	236	2	Q6ZAE9_ORYSA	Q6zae9 oryza sativ
459	36	64.3	1060	2	Q5QGH9_S1VCZ	Q5qgh9 chimpanzee	532	35	62.5	239	2	Q7W1D9_BORPA	Q7w1d9 bordetella
460	36	64.3	1060	2	Q5QGH0_S1VCZ	Q5qgh0 chimpanzee	533	35	62.5	244	2	Q7W0F8_BORPE	Q7w0f8 bordetella
461	36	64.3	1060	2	Q5QGF7_S1VCZ	Q5qgf7 chimpanzee	534	35	62.5	244	2	Q7WP46_BORBR	Q7wp46 bordetella
462	36	64.3	1060	2	Q5QGB6_S1VCZ	Q5qgb6 chimpanzee	535	35	62.5	246	2	Q8N3X4_HUMAN	Q8n3x4 homo sapien
463	36	64.3	1060	2	Q5QGA9_S1VCZ	Q5qga9 chimpanzee	536	35	62.5	249	2	Q658K5_MANSN	Q658k5 manneheimia
464	36	64.3	1060	2	Q5QG99_S1VCZ	Q5qg99 chimpanzee	537	35	62.5	257	2	Q8ELB3_OCETH	Q8elb3 oceanobacil
465	36	64.3	1060	2	Q5QG92_S1VCZ	Q5qg92 chimpanzee	538	35	62.5	281	2	Q8Y2V7_RALSO	Q8y2v7 ralstonia s
466	36	64.3	1060	2	Q5QGX8_S1VCZ	Q5qgx8 chimpanzee	539	35	62.5	291	2	Q7M7W7_WOLNELA	Q7m7w7 wolnella s
467	36	64.3	1060	2	Q5QFX8_S1VCZ	Q5qfx8 chimpanzee	540	35	62.5	292	2	Q5GYG8_XANOR	Q5gyg8 xanthomonas
468	36	64.3	1060	2	Q5QFV7_S1VCZ	Q5qfv7 chimpanzee	541	35	62.5	293	2	Q4J489_AZOVI	Q4j489 azotobacter
469	36	64.3	1060	2	Q5QFU2_S1VCZ	Q5qfu2 chimpanzee	542	35	62.5	297	1	XERC_MYCLE	Q9cbu0 mycobacteri

543	35	62.5	305	2	Q7NRM4	CHRV0	Q7nrm4 chromobacte	616	35	62.5	577	2	Q7NL26_GLOVI	Q7nl26 gloebacter
544	35	62.5	306	2	QAOIH6_LEIMA	Qagih6 leishmania	617	35	62.5	577	2	Q9IN28_9HIV1	Q9in28 human immun	
545	35	62.5	307	1	PCPR_SPICR	P26f79 sphingobium	618	35	62.5	582	2	Q8AB47_BACTN	Q8ab47 bacteroides	
546	35	62.5	308	2	QXLL16_AERHY	Qxll16 aeromonas h	619	35	62.5	598	2	Q87Q08_VIBPA	Q87q08 vibrio para	
547	35	62.5	317	2	Q8H479_ORYSA	Q8h479 oryza sativ	620	35	62.5	610	2	Q696X6_ARATH	Q696x6 arabidopsis	
548	35	62.5	323	2	Q6AGE3_LEIXX	Q6age3 leifsonia x	621	35	62.5	610	2	Q696X8_ARATH	Q696x8 arabidopsis	
549	35	62.5	328	2	Q5V5P2_HALMA	Q5v5f2 haloarcula	622	35	62.5	610	2	Q69627_ARATH	Q69627 arabidopsis	
550	35	62.5	334	2	Q5OZ31_IDILO	Q5qz31 idiomarina	623	35	62.5	610	2	Q697A2_ARATH	Q697a2 arabidopsis	
551	35	62.5	335	2	Q50A80_9HIV1	Q50a80 human immun	624	35	62.5	612	1	CRY2_ARATH	Q96524 arabidopsis	
552	35	62.5	342	2	Q89D19_BRAJA	Q89d19 bradyrhizob	625	35	62.5	617	2	Q84L82_ARMRU	Q84l82 armoracia r	
553	35	62.5	343	2	Q4LX28_9EURK	Q4lx28 burkholderi	626	35	62.5	617	2	Q84L83_ARMRU	Q84l83 armoracia r	
554	35	62.5	349	2	Q5R0S2_IDILO	Q5r0s2 idiomarina	627	35	62.5	617	2	Q84L84_ARMRU	Q84l84 armoracia r	
555	35	62.5	351	2	Q72AN6_DESVH	Q72an6 desulfovibr	628	35	62.5	617	2	Q84UJ4_ARMRU	Q84uj4 armoracia r	
556	35	62.5	353	2	Q5AY13_EMENI	Q5ay13 aspergillus	629	35	62.5	619	2	Q4KXK0_HUMAN	Q4kxk0 homo sapien	
557	35	62.5	356	2	Q50HW3_9HIV1	Q50hw3 human immun	630	35	62.5	631	2	Q5FT72_GLUOX	Q5ft72 gluconobact	
558	35	62.5	359	2	Q5V5G5_HALMA	Q5v5g5 haloarcula	631	35	62.5	635	2	Q9A882_CAUCR	Q9a882 caulobacter	
559	35	62.5	370	1	CYB_MICIK	Q9mlk2 micropechis	632	35	62.5	647	2	Q6FV06_CANGA	Q6fv06 candida gla	
560	35	62.5	370	2	Q6N0D3_RHOPA	Q6n0d3 rhodopseudo	633	35	62.5	653	2	Q89MC3_BRAJA	Q89mc3 bradyrhizob	
561	35	62.5	372	1	3BHS_PIG	Q9h119 s 3 beta-hy	634	35	62.5	663	2	Q81YU8_HUMAN	Q81y8 homo sapien	
562	35	62.5	372	2	Q5SJG6_THET8	Q5sjw6 thermus the	635	35	62.5	669	2	Q8DKV3_SYNEL	Q8dkv3 synechococc	
563	35	62.5	372	2	Q72KB8_THET2	Q72kb8 thermus the	636	35	62.5	679	2	Q51U94_MAGGR	Q51u94 magnaporth	
564	35	62.5	379	2	Q8DM04_SYNEL	Q8dm04 synechococc	637	35	62.5	690	1	EPB42_MOUSE	P49222 mus musculu	
565	35	62.5	381	1	CYB_PHADO	Q9rd01 streptomyce	638	35	62.5	690	1	EPB42_MOUSE	Q4vb97 homo sapien	
566	35	62.5	381	2	Q5JLU0_PHADO	Q5j425 phascosolor	639	35	62.5	691	2	Q4VB97_HUMAN	Q4vb97 homo sapien	
567	35	62.5	383	2	Q4KH3_PSEF5	Q4khu0 phascosolor	640	35	62.5	691	2	Q6NXZ8_MOUSE	Q6nxz8 mus musculu	
568	35	62.5	395	2	Q5P6Q0_AZOSE	Q5p6q0 azocarcus sp	641	35	62.5	705	2	Q53SX5_HUMAN	Q53sx5 homo sapien	
569	35	62.5	410	2	Q7WS65_AUCEU	Q7ws65 aicaligenes	642	35	62.5	710	2	Q8RY11_ARATH	Q8ry11 arabidopsis	
570	35	62.5	415	2	Q52212_9ZZZZ	Q52212 plasmid pae	643	35	62.5	712	2	Q7MU14_PORGI	Q7mu14 porphyromon	
571	35	62.5	416	2	Q8FPM9_COREF	Q8fpm9 corynebacte	644	35	62.5	722	2	Q6ND23_ERWCT	Q6nd23 erwinia car	
572	35	62.5	419	2	Q5B6X9_EMENI	Q5b6x9 aspergillus	645	35	62.5	746	2	Q5FSY4_GLUOX	Q5fsy4 gluconobact	
573	35	62.5	420	2	Q8Q739_9HIV1	Q8q739 human immun	646	35	62.5	758	2	Q8N9H8_HUMAN	Q8n9h8 homo sapien	
574	35	62.5	423	2	Q8QDC1_9HIV1	Q8qdc1 human immun	647	35	62.5	760	2	Q5GZRA_XANOR	Q5gzra xanthomonas	
575	35	62.5	433	2	Q9RD01_STRCO	Q9rd01 streptomyce	648	35	62.5	760	2	Q4UUK0_XANCP	Q4uuk0 xanthomonas	
576	35	62.5	434	2	Q6JUN9_9HIV1	Q6j1j9 human immun	649	35	62.5	760	2	Q8P998_XANCP	Q8p998 xanthomonas	
577	35	62.5	434	2	Q6JUN2_9HIV1	Q6j1j2 human immun	650	35	62.5	760	2	Q8P998_XANCP	Q8p998 xanthomonas	
578	35	62.5	434	2	Q6JJJ9_9HIV1	Q6j1j9 human immun	651	35	62.5	766	2	Q8PL05_XANCP	Q8pl05 xanthomonas	
579	35	62.5	434	2	Q6H1W7_9HIV1	Q6h1w7 human immun	652	35	62.5	791	2	Q9WZ02_PSEBH	Q9wz02 pseudomonas	
580	35	62.5	434	2	Q5G532_9HIV1	Q5g532 human immun	653	35	62.5	793	1	Q9KU26_VIBCH	Q9ku26 vibrio chol	
581	35	62.5	435	2	Q5LS05_MAGGR	Q5ls05 magnaporth	654	35	62.5	793	1	K1F3C_HUMAN	K1f3c human	
582	35	62.5	437	2	Q5AMQ1_CANAL	Q5amq1 candida alb	655	35	62.5	793	2	Q722W5_HUMAN	Q722w5 homo sapien	
583	35	62.5	443	2	Q7R3Z8_GIALA	Q7r3z8 giardia lam	656	35	62.5	795	2	Q562P7_HUMAN	Q562p7 homo sapien	
584	35	62.5	446	2	Q5SLH9_CRYNE	Q5slh9 cryptococcu	657	35	62.5	795	2	Q53781_MYCTU	Q53781 mycobacteri	
585	35	62.5	446	2	Q5KA18_CRYNE	Q5kal8 cryptococcu	658	35	62.5	795	2	Q7U1Q0_MYCBO	Q7u1q0 mycobacteri	
586	35	62.5	446	2	Q4NNSC_9BELT	Q4nns8 anaeromyxob	659	35	62.5	803	2	Q9KKW3_VIBCH	Q9kkw3 vibrio chol	
587	35	62.5	451	1	Y4337_RHIME	Q33683 rhizobium m	660	35	62.5	819	2	Q5H1C4_XANOR	Q5h1c4 xanthomonas	
588	35	62.5	454	2	Q8DGB6_SYNEL	Q8dgb6 synechococc	661	35	62.5	848	2	Q4RVV3_TFTNG	Q4rvv3 tetraodon n	
589	35	62.5	467	2	Q6NSQ8_MOUSE	Q6nsq8 mus musculu	662	35	62.5	859	2	Q7MTD8_PORGI	Q7mtd8 porphyromon	
590	35	62.5	472	1	PHR_ECOLI	P00914 escherichia	663	35	62.5	904	1	DPO1_MYCBO	P0a551 mycobacteri	
591	35	62.5	472	2	Q7AGL3_ECO57	Q7agl3 escherichia	664	35	62.5	904	1	DPO1_MYCBO	P0a550 mycobacteri	
592	35	62.5	472	2	Q8FJV0_ECOL6	Q8fjv0 escherichia	665	35	62.5	915	2	Q6P179_HUMAN	Q6p179 homo sapien	
593	35	62.5	472	2	Q8X9D4_ECO57	Q8x9d4 escherichia	666	35	62.5	915	2	Q7XSI0_ORYSA	Q7xsi0 oryza sativ	
594	35	62.5	475	2	Q52QJ0_PRODI	Q52qj0 prochloron	667	35	62.5	925	2	Q7XSI0_ORYSA	Q7xsi0 oryza sativ	
595	35	62.5	480	2	Q22813_ARATH	Q22813 arabidopsis	668	35	62.5	931	2	Q7WY21_PSEAE	Q7wy21 pseudomonas	
596	35	62.5	486	2	Q9SQ00_ARATH	Q9sq00 arabidopsis	669	35	62.5	931	2	Q7WY21_PSEAE	Q7wy21 pseudomonas	
597	35	62.5	488	2	Q93383_CABEL	Q93383 caenorhabdi	670	35	62.5	933	2	Q677U1_VIRU	Q677u1 lymphocyti	
598	35	62.5	489	2	Q62100_CABEL	Q62100 caenorhabdi	671	35	62.5	960	2	Q7Z5K1_HUMAN	Q7z5k1 homo sapien	
599	35	62.5	491	2	Q7U9L8_SYNPX	Q7u9l8 synechococc	672	35	62.5	960	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
600	35	62.5	501	1	PHR1_SINAL	P40115 sinapis alb	673	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
601	35	62.5	508	2	Q6GNG6_XENLA	Q6gng6 xenopus lae	674	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
602	35	62.5	519	2	Q8A7M4_BACTN	Q8a7m4 bacteroides	675	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
603	35	62.5	521	2	Q82507_ARATH	Q82507 arabidopsis	676	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
604	35	62.5	522	2	Q5BGN6_EMENI	Q5bgng6 aspergillus	677	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
605	35	62.5	522	2	Q8VKI3_MYCTU	Q8vki3 mycobacteri	678	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
606	35	62.5	524	2	Q9ZKU3_HELPJ	Q9zku3 helicobacte	679	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
607	35	62.5	526	2	Q4V789_XENTR	Q4v789 xenopus tro	680	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
608	35	62.5	527	2	Q25564_HELPY	Q25564 helicobacte	681	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
609	35	62.5	534	2	Q5ZSM0_9HIV2	Q5zsm0 human immun	682	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
610	35	62.5	542	2	Q6APS4_DESPS	Q6ape4 desulfotale	683	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
611	35	62.5	553	2	Q7WGH7_BORBR	Q7wgh7 bordetella	684	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
612	35	62.5	553	2	Q7WVX5_BORPA	Q7wv12 bordetella	685	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
613	35	62.5	553	2	Q7VVKX5_BORPE	Q7vvx5 bordetella	686	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
614	35	62.5	558	2	Q5JMX2_ORYSA	Q5jmx2 oryza sativ	687	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	
615	35	62.5	569	2	Q9MA84_ARATH	Q9ma84 arabidopsis	688	35	62.5	965	2	Q5RFP3_PONPY	Q5rfp3 pongo pygma	

689	35	62.5	1722	2	Q53L67_ORYSA	Q53L67 oryza sativ	762	34	60.7	178	2	Q9B595_CHILA	Q9B595 chinchilla
690	35	62.5	2111	2	Q5AV07_EMENI	Q5av07 aspergillus	763	34	60.7	179	2	Q94Q16_CHIBR	Q94Q16 chinchilla
691	35	62.5	2229	2	Q9EWP5_STRCO	Q9ewp5 streptomyce	764	34	60.7	182	2	Q94Q13_CHILA	Q94Q13 chinchilla
692	35	62.5	3433	1	PCX_DRONE	P18490 drosophila	765	34	60.7	182	2	Q94Q14_CHILA	Q94Q14 chinchilla
693	35	62.5	5072	2	Q6JD66_TOXGO	Q6jd66 toxoplasma	766	34	60.7	182	2	Q94Q12_CHIBR	Q94Q12 chinchilla
694	35	62.5	7599	2	Q4IK83_GIBZE	Q4ik83 gibberella	767	34	60.7	182	2	Q9B592_CHIBR	Q9B592 chinchilla
695	34.5	61.6	294	2	Q4HC20_DEIO	Q4hc20 deinococcus	768	34	60.7	182	2	Q8HH86_CHILA	Q8HH86 chinchilla
696	34	60.7	32	2	Q4VUT0_STRPY	Q4vut0 streptococc	769	34	60.7	182	2	Q4VTW0_9SAUR	Q4vtw0 gymmodactyl
697	34	60.7	38	2	Q7YGP4_9HYST	Q7ygp4 ctenomys so	770	34	60.7	183	2	Q999D3_9HIV1	Q999D3 human immun
698	34	60.7	70	1	RK28_CYAPA	P48129 cyanophora	771	34	60.7	184	2	Q912L8_PSAE	Q912L8 pseudomonas
699	34	60.7	71	2	Q64GL7_9HIV1	Q64gl7 human immun	772	34	60.7	185	2	Q4VTV4_9SAUR	Q4vtv4 gymmodactyl
700	34	60.7	85	2	Q9R835_DEIRA	Q9r835 deinococcus	773	34	60.7	186	2	Q8HH87_CHILA	Q8HH87 chinchilla
701	34	60.7	85	2	Q64GN1_9HIV1	Q64gn1 human immun	774	34	60.7	186	2	Q8HH85_CHILA	Q8HH85 chinchilla
702	34	60.7	93	2	Q71E12_BACSU	Q71e12 bacillus su	775	34	60.7	186	2	Q8HBP8_CHIBR	Q8HBP8 chinchilla
703	34	60.7	97	2	Q8FPC4_COREF	Q8fpc4 corynebacte	776	34	60.7	187	2	Q4VTY4_9SAUR	Q4vty4 gymmodactyl
704	34	60.7	108	2	Q8E992_SHEON	Q8e992 shewanella	777	34	60.7	187	2	Q4VTX9_9SAUR	Q4vtx9 gymmodactyl
705	34	60.7	115	2	Q98HW9_RHILO	Q98hw9 rhizobium l	778	34	60.7	190	2	Q4VTV1_9SAUR	Q4vtv1 gymmodactyl
706	34	60.7	118	2	Q6UIA9_9HIV1	Q6uia9 human immun	779	34	60.7	191	2	Q98ZG9_9HIV1	Q98zg9 human immun
707	34	60.7	127	2	Q97468_9HIV1	Q97468 human immun	780	34	60.7	192	2	Q6N4D2_RHOPA	Q6n4d2 rhodopseudo
708	34	60.7	127	2	Q97437_9HIV1	Q97437 human immun	781	34	60.7	192	2	Q6X4W7_9HYST	Q6x4w7 lagidium vi
709	34	60.7	127	2	Q97443_9HIV1	Q97443 human immun	782	34	60.7	192	2	Q6X4W6_CHILA	Q6x4w6 chinchilla
710	34	60.7	127	2	Q97480_9HIV1	Q97480 human immun	783	34	60.7	193	2	Q7NV13_CHRVO	Q7nv13 chromobacte
711	34	60.7	127	2	Q97441_9HIV1	Q97441 human immun	784	34	60.7	193	2	Q6X4W6_9HYST	Q6x4w6 lagidium vi
712	34	60.7	127	2	Q97463_9HIV1	Q97463 human immun	785	34	60.7	194	2	Q4VTW2_9SAUR	Q4vtw2 gymmodactyl
713	34	60.7	127	2	Q97430_9HIV1	Q97430 human immun	786	34	60.7	197	2	Q4VTX0_9SAUR	Q4vtx0 gymmodactyl
714	34	60.7	127	2	Q97428_9HIV1	Q97428 human immun	787	34	60.7	198	2	Q4VTX6_9SAUR	Q4vtx6 gymmodactyl
715	34	60.7	127	2	Q97434_9HIV1	Q97434 human immun	788	34	60.7	198	2	Q4VTV6_9SAUR	Q4vtv6 gymmodactyl
716	34	60.7	127	2	Q97461_9HIV1	Q97461 human immun	789	34	60.7	201	2	Q4ZXP3_PSESY	Q4zxp3 pseudomonas
717	34	60.7	127	2	Q97459_9HIV1	Q97459 human immun	790	34	60.7	201	2	Q887U8_PSESM	Q887u8 pseudomonas
718	34	60.7	127	2	Q97439_9HIV1	Q97439 human immun	791	34	60.7	203	2	Q6N8P8_RHOPA	Q6n8p8 rhodopseudo
719	34	60.7	127	2	Q97457_9HIV1	Q97457 human immun	792	34	60.7	204	2	Q4VTY3_9SAUR	Q4vty3 gymmodactyl
720	34	60.7	127	2	Q97432_9HIV1	Q97432 human immun	793	34	60.7	206	2	Q8Y173_RALSO	Q8y173 ralstonia s
721	34	60.7	127	2	Q97455_9HIV1	Q97455 human immun	794	34	60.7	206	2	Q6Q6C0_9HIV1	Q6q6c0 human immun
722	34	60.7	133	1	RL22_AQUAE	Q66436 aquifex aeo	795	34	60.7	206	2	Q6Q668_9HIV1	Q6q668 human immun
723	34	60.7	139	2	Q9I225_9HIV1	Q9i225 human immun	796	34	60.7	207	2	Q6X4W9_9HYST	Q6x4w9 lagidium pe
724	34	60.7	144	2	Q7XWC2_ORYSA	Q7xwc2 oryza sativ	797	34	60.7	207	2	Q6X4W8_9HYST	Q6x4w8 lagidium vi
725	34	60.7	155	2	Q5DSV7_CUCUC	Q5dsv7 caryedon se	798	34	60.7	207	2	Q6X4W5_9HYST	Q6x4w5 lagidium vi
726	34	60.7	158	1	YBJN_ECOLI	P75815 escherichia	799	34	60.7	207	2	Q6Q658_9HIV1	Q6q658 human immun
727	34	60.7	158	2	Q6S2U6_ORYSA	Q6s2u6 oryza sativ	800	34	60.7	207	2	Q6Q657_9HIV1	Q6q657 human immun
728	34	60.7	158	2	Q57R86_SALCH	Q57r86 salmonella	801	34	60.7	207	2	Q6Q642_9HIV1	Q6q642 human immun
729	34	60.7	158	2	Q5PGM6_SALPA	Q5pgm6 salmonella	802	34	60.7	207	2	Q6Q641_9HIV1	Q6q641 human immun
730	34	60.7	158	2	Q8Z847_SALTI	Q8z847 salmonella	803	34	60.7	207	2	Q6Q640_9HIV1	Q6q640 human immun
731	34	60.7	158	2	Q8ZOK1_SALTY	Q8zok1 salmonella	804	34	60.7	209	2	Q9L059_STRCO	Q9l059 streptomyce
732	34	60.7	160	2	Q742N2_MYCPA	Q742n2 mycobacteri	805	34	60.7	209	2	Q92PA0_RHIME	Q92pa0 rhizobium m
733	34	60.7	160	2	Q83DU0_COXBU	Q83du0 coxiella bu	806	34	60.7	210	2	Q7XFG5_ORYSA	Q7xfg5 oryza sativ
734	34	60.7	164	2	Q9B597_CHILA	Q9b597 chinchilla	807	34	60.7	210	2	Q4M113_9BURK	Q4m113 burkholderi
735	34	60.7	166	2	Q67F75_9HIV1	Q67f75 human immun	808	34	60.7	210	2	Q4MTM6_9SAUR	Q4mtm6 gymmodactyl
736	34	60.7	166	2	Q67F75_9HIV1	Q67f75 human immun	809	34	60.7	211	2	Q98Y80_9HIV1	Q98y80 human immun
737	34	60.7	166	2	Q67I08_9HIV1	Q67i08 human immun	810	34	60.7	211	2	Q9ECL6_9HIV1	Q9ec16 human immun
738	34	60.7	166	2	Q67I18_9HIV1	Q67i18 human immun	811	34	60.7	213	2	Q6D7L9_ERWCT	Q6d7l9 erwina car
739	34	60.7	166	2	Q67I17_9HIV1	Q67i17 human immun	812	34	60.7	214	2	Q63IV1_BURPS	Q63iv1 burkholderi
740	34	60.7	166	2	Q67I08_9HIV1	Q67i08 human immun	813	34	60.7	214	2	Q62ED0_BURMA	Q62ed0 burkholderi
741	34	60.7	166	2	Q67I07_9HIV1	Q67i07 human immun	814	34	60.7	215	2	Q6UZR2_9HIV1	Q6uzr2 human immun
742	34	60.7	166	2	Q67I04_9HIV1	Q67i04 human immun	815	34	60.7	217	2	Q4P6B1_USTMA	Q4p6b1 ustilago ma
743	34	60.7	166	2	Q67I02_9HIV1	Q67i02 human immun	816	34	60.7	218	2	Q5AZJ2_EMENI	Q5azj2 aspergillus
744	34	60.7	166	2	Q67I01_9HIV1	Q67i01 human immun	817	34	60.7	218	2	Q4VTV8_9SAUR	Q4vtv8 gymmodactyl
745	34	60.7	166	2	Q67I08_9HIV1	Q67i08 human immun	818	34	60.7	218	2	Q5MI28_9HIV1	Q5mi28 human immun
746	34	60.7	166	2	Q67I07_9HIV1	Q67i07 human immun	819	34	60.7	219	2	Q89943_9HIV1	Q89943 human immun
747	34	60.7	166	2	Q67I08_9HIV1	Q67i08 human immun	820	34	60.7	219	2	Q9ID86_9HIV1	Q9id86 human immun
748	34	60.7	166	2	Q67I09_9HIV1	Q67i09 human immun	821	34	60.7	220	1	NADD_YERPE	Q8zdg1 yersinia pe
749	34	60.7	166	2	Q67I03_9HIV1	Q67i03 human immun	822	34	60.7	220	2	Q6DDE6_YERPS	Q6dde6 yersinia ps
750	34	60.7	166	2	Q67I03_9HIV1	Q67i03 human immun	823	34	60.7	220	2	Q8SH40_9SAUR	Q8sh40 pachyactyl
751	34	60.7	166	2	Q67I37_9HIV1	Q67i37 human immun	824	34	60.7	220	2	Q4VTV5_9SAUR	Q4vtv5 gymmodactyl
752	34	60.7	166	2	Q67I21_9HIV1	Q67i21 human immun	825	34	60.7	221	2	Q9V7I9_LACAG	Q9v7i9 lacerta agi
753	34	60.7	166	2	Q67I20_9HIV1	Q67i20 human immun	826	34	60.7	221	2	Q8M2E4_9SAUR	Q8m2e4 pachyactyl
754	34	60.7	167	2	Q9B599_CHILA	Q9b599 chinchilla	827	34	60.7	221	2	Q8M2E1_9SAUR	Q8m2e1 pachyactyl
755	34	60.7	168	2	Q4TP22_TETNG	Q4tp22 tetradodon n	828	34	60.7	221	2	Q8M2D2_9SAUR	Q8m2d2 pachyactyl
756	34	60.7	169	2	Q94Q11_9HYST	Q94q11 octomys mim	829	34	60.7	221	2	Q8M2E2_9SAUR	Q8m2e2 pachyactyl
757	34	60.7	169	2	Q9ECC2_9HIV1	Q9ecc2 human immun	830	34	60.7	221	2	Q5MI49_9HIV1	Q5mi49 human immun
758	34	60.7	169	2	Q9QF21_9HIV1	Q9qf21 human immun	831	34	60.7	222	2	Q9ID73_9HIV1	Q9id73 human immun
759	34	60.7	170	2	Q9B5A0_CHILA	Q9b5a0 chinchilla	832	34	60.7	222	2	Q5RM10_9HIV1	Q5rm10 human immun
760	34	60.7	171	2	Q8Q8W3_9HIV1	Q8q8w3 human immun	833	34	60.7	223	2	Q4VTW7_9SAUR	Q4vtw7 gymmodactyl
761	34	60.7	178	2	Q84X11_LYCES	Q84x11 lycopersico	834	34	60.7	223	2	Q4VTW4_9SAUR	Q4vtw4 gymmodactyl

835	34	60.7	223	2	Q999D2_9HIV1	Q999d2 human immun	908	34	60.7	238	2	Q5MEW4_9HIV1	Q5mhv4 human immun
836	34	60.7	224	2	Q98X05_9HIV1	Q98x05 human immun	909	34	60.7	238	2	Q5O9Q1_9HIV1	Q5O9q1 human immun
837	34	60.7	224	2	Q98MW5_9HIV1	Q98mw5 human immun	910	34	60.7	239	2	Q7SM78_9HIV1	Q7sm78 human immun
838	34	60.7	224	2	Q98WZ9_9HIV1	Q98wz9 human immun	911	34	60.7	239	2	Q6W37_9HIV1	Q6w37 human immun
839	34	60.7	224	2	Q98X19_9HIV1	Q98x19 human immun	912	34	60.7	240	2	Q7ZKP5_9HIV1	Q7zkp5 human immun
840	34	60.7	224	2	Q98W21_9HIV1	Q98wz1 human immun	913	34	60.7	241	2	Q90M39_9HIV1	Q90m39 human immun
841	34	60.7	224	2	Q99EU4_9HIV1	Q99eu4 human immun	914	34	60.7	241	2	Q5RM97_9HIV1	Q5rm97 human immun
842	34	60.7	224	2	Q98X02_9HIV1	Q98x02 human immun	915	34	60.7	242	2	Q9T7E0_9HIV1	Q9t7e0 trinomys gr
843	34	60.7	224	2	Q5MHX2_9HIV1	Q5mhx2 human immun	916	34	60.7	242	2	Q9T7D9_9HIV1	Q9t7d9 trinomys gr
844	34	60.7	225	2	Q4VTV9_9SAUR	Q4vtv9 gymnodactyl	917	34	60.7	242	2	Q9T7E1_9HIV1	Q9t7e1 trinomys gr
845	34	60.7	225	2	Q999G4_9HIV1	Q999g4 human immun	918	34	60.7	242	2	Q8SPI7_9HIV1	Q8spi7 sigmodon hi
846	34	60.7	226	2	Q4VTW9_9SAUR	Q4vtw9 gymnodactyl	919	34	60.7	242	2	Q8SELO_9RODE	Q8sel0 sigmodon ma
847	34	60.7	226	2	Q8Q3T7_9HIV1	Q8q3t7 human immun	920	34	60.7	242	2	Q9T4B5_9HIV1	Q9t4b5 trinomys gr
848	34	60.7	226	2	Q99ET8_9HIV1	Q99et8 human immun	921	34	60.7	242	2	Q4VTX4_9SAUR	Q4vtx4 gymnodactyl
849	34	60.7	226	2	Q99ET7_9HIV1	Q99et7 human immun	922	34	60.7	243	2	Q5MI35_9HIV1	Q5mi35 human immun
850	34	60.7	226	2	Q99EV4_9HIV1	Q99ev4 human immun	923	34	60.7	243	2	Q5MHT7_9HIV1	Q5mht7 human immun
851	34	60.7	226	2	Q99D46_9HIV1	Q99d46 human immun	924	34	60.7	244	2	Q5B7M7_9HIV1	Q5b7m7 aspergillus
852	34	60.7	226	2	Q66W22_9HIV1	Q66w22 human immun	925	34	60.7	244	2	Q90BB3_9HIV1	Q90bb3 human immun
853	34	60.7	227	2	Q4VTV7_9SAUR	Q4vtv7 gymnodactyl	926	34	60.7	244	2	Q7ZKJ9_9HIV1	Q7zjk9 human immun
854	34	60.7	227	2	Q90RM9_9HIV1	Q90rm9 human immun	927	34	60.7	244	2	Q7ZKJ4_9HIV1	Q7zjk4 human immun
855	34	60.7	227	2	Q992D3_9HIV1	Q992d3 human immun	928	34	60.7	244	2	Q5MI44_9HIV1	Q5mi44 human immun
856	34	60.7	227	2	Q992B3_9HIV1	Q992b3 human immun	929	34	60.7	244	2	Q5MI37_9HIV1	Q5mi37 human immun
857	34	60.7	227	2	Q5MHV6_9HIV1	Q5mhv6 human immun	930	34	60.7	244	2	Q5MI27_9HIV1	Q5mi27 human immun
858	34	60.7	228	2	Q9MHR7_9SAUR	Q9mhr7 pachydactyl	931	34	60.7	244	2	Q5MHU6_9HIV1	Q5mhu6 human immun
859	34	60.7	228	2	Q4VTW5_9SAUR	Q4vtw5 gymnodactyl	932	34	60.7	245	2	Q7ZKM6_9HIV1	Q7zkm6 human immun
860	34	60.7	229	2	Q70EG3_9RODE	Q70eg3 saccostomus	933	34	60.7	245	2	Q7ZKI3_9HIV1	Q7zki3 human immun
861	34	60.7	229	2	Q9Q302_9HIV1	Q9q302 human immun	934	34	60.7	245	2	Q7ZFK0_9HIV1	Q7zfk0 human immun
862	34	60.7	229	2	Q9Q380_9HIV1	Q9q380 human immun	935	34	60.7	245	2	Q7ZFK6_9HIV1	Q7zfk6 human immun
863	34	60.7	229	2	Q9Q316_9HIV1	Q9q316 human immun	936	34	60.7	245	2	Q6QJUL_9HIV1	Q6qjul human immun
864	34	60.7	229	2	Q5MHV8_9HIV1	Q5mhv8 human immun	937	34	60.7	245	2	Q5MI16_9HIV1	Q5mi16 human immun
865	34	60.7	229	2	Q50BA3_9HIV1	Q5oba3 human immun	938	34	60.7	246	2	Q7ZKL6_9HIV1	Q7zkl6 human immun
866	34	60.7	230	2	Q4VTX1_9SAUR	Q4vtx1 gymnodactyl	939	34	60.7	247	2	Q7ZKE0_9HIV1	Q7zke0 human immun
867	34	60.7	230	2	Q4VTU9_9SAUR	Q4vtu9 gymnodactyl	940	34	60.7	247	2	Q5MI50_9HIV1	Q5mi50 human immun
868	34	60.7	230	2	Q9Q392_9HIV1	Q9q392 human immun	941	34	60.7	247	2	Q5MI06_9HIV1	Q5mi06 human immun
869	34	60.7	230	2	Q992E6_9HIV1	Q992e6 human immun	942	34	60.7	247	2	Q5MHV0_9HIV1	Q5mhv0 human immun
870	34	60.7	230	2	Q5MI09_9HIV1	Q5mi09 human immun	943	34	60.7	249	2	Q5MHU2_9HIV1	Q5mhu2 human immun
871	34	60.7	231	2	Q4VTW3_9SAUR	Q4vtw3 gymnodactyl	944	34	60.7	249	2	Q7ZKG2_9HIV1	Q7zkg2 human immun
872	34	60.7	232	2	Q6N2F2_9ROPA	Q6n2f2 rhodopseudo	945	34	60.7	249	2	Q5RSY6_9HIV1	Q5rsy6 human immun
873	34	60.7	232	2	Q4VTX8_9SAUR	Q4vtx8 gymnodactyl	946	34	60.7	249	2	Q5RSY5_9HIV1	Q5rsy5 human immun
874	34	60.7	232	2	Q4VTW1_9SAUR	Q4vtw1 gymnodactyl	947	34	60.7	249	2	Q5RR44_9HIV1	Q5rr44 human immun
875	34	60.7	232	2	Q999Q8_9HIV1	Q999q8 human immun	948	34	60.7	250	2	Q5RQ83_9HIV1	Q5rq83 human immun
876	34	60.7	232	2	Q85Q52_9HIV1	Q85q52 trichomys	949	34	60.7	250	2	Q90M40_9HIV1	Q90m40 human immun
877	34	60.7	233	2	Q4VTY2_9SAUR	Q4vtv2 gymnodactyl	950	34	60.7	250	2	Q90M44_9HIV1	Q90m44 human immun
878	34	60.7	233	2	Q4VTW8_9SAUR	Q4vtw8 gymnodactyl	951	34	60.7	250	2	Q71309_9HIV1	Q71309 human immun
879	34	60.7	233	2	Q4VTW3_9SAUR	Q4vtw3 gymnodactyl	952	34	60.7	250	2	Q991W5_9HIV1	Q991w5 human immun
880	34	60.7	233	2	Q6QJRS_9HIV1	Q6qjrs human immun	953	34	60.7	250	2	Q71308_9HIV1	Q71308 human immun
881	34	60.7	233	2	Q5MHX5_9HIV1	Q5mhx5 human immun	954	34	60.7	250	2	Q71310_9HIV1	Q71310 human immun
882	34	60.7	234	2	Q4K3V9_PSEF5	Q4k3v9 pseudomonas	955	34	60.7	250	2	Q6QJR6_9HIV1	Q6qjr6 human immun
883	34	60.7	234	2	Q4VTX7_9SAUR	Q4vtx7 gymnodactyl	956	34	60.7	251	2	Q5HLW9_STAEP	Q5hlw9 staphylococ
884	34	60.7	234	2	Q999E7_9HIV1	Q999e7 human immun	957	34	60.7	251	2	Q8CRB9_STAEP	Q8cre9 staphylococ
885	34	60.7	234	2	Q50B99_9HIV1	Q50b99 human immun	958	34	60.7	251	2	Q5RNL8_9HIV1	Q5rnl8 human immun
886	34	60.7	234	2	Q50B86_9HIV1	Q50b86 human immun	959	34	60.7	252	2	Q8PT86_COREF	Q8ft86 corynebacte
887	34	60.7	234	2	Q50B27_9HIV1	Q50b27 human immun	960	34	60.7	252	2	Q6NHS3_CORGL	Q6nhq7 corynebacte
888	34	60.7	234	2	Q509V2_9HIV1	Q509v2 human immun	961	34	60.7	252	2	Q8NQ77_CORGL	Q8nq77 corynebacte
889	34	60.7	235	2	Q4VTY1_9SAUR	Q4vtv1 gymnodactyl	962	34	60.7	252	2	Q4VTV2_9SAUR	Q4vtv2 gymnodactyl
890	34	60.7	235	2	Q4VTX5_9SAUR	Q4vtv5 gymnodactyl	963	34	60.7	253	2	Q6N745_RHOPA	Q6n745 rhodopseudo
891	34	60.7	235	2	Q5MHW8_9HIV1	Q5mhw8 human immun	964	34	60.7	253	2	Q90CM2_9HIV1	Q90cm2 human immun
892	34	60.7	235	2	Q50B76_9HIV1	Q50b76 human immun	965	34	60.7	253	2	Q7ZKL3_9HIV1	Q7zkl3 human immun
893	34	60.7	235	2	Q53A80_9HIV1	Q53a80 human immun	966	34	60.7	254	2	Q5RR25_9HIV1	Q5rr25 human immun
894	34	60.7	236	2	Q6XNT0_9SAUR	Q6xnt0 phlesuma gu	967	34	60.7	254	2	Q6TV17_9HIV1	Q6tcv17 isothrix bi
895	34	60.7	237	2	Q7ZKD7_9HIV1	Q7zkd7 human immun	968	34	60.7	255	2	Q7ZKK6_9HIV1	Q7zkk6 human immun
896	34	60.7	237	2	Q68Q78_9HIV1	Q68q78 human immun	969	34	60.7	255	2	Q7ZKE7_9HIV1	Q7zke7 human immun
897	34	60.7	237	2	Q6XNZ7_9SAUR	Q6xnz7 pachydactyl	970	34	60.7	255	2	Q7ZKF4_9HIV1	Q7zfk4 human immun
898	34	60.7	238	2	Q6XNZ3_9SAUR	Q6xnz3 phlesuma ma	971	34	60.7	255	2	Q6QTR0_9HIV1	Q6qtr0 human immun
899	34	60.7	238	2	Q6XNY9_9SAUR	Q6xny9 phlesuma mu	972	34	60.7	255	2	Q6QTP2_9HIV1	Q6qtp2 human immun
900	34	60.7	238	2	Q6XNY6_9SAUR	Q6xny6 phlesuma li	973	34	60.7	255	2	Q6QTP1_9HIV1	Q6qtp1 human immun
901	34	60.7	238	2	Q6XNY3_9SAUR	Q6xny3 phlesuma qu	974	34	60.7	255	2	Q5RRL7_9HIV1	Q5rrl7 human immun
902	34	60.7	238	2	Q6XNY2_9SAUR	Q6xny2 phlesuma ab	975	34	60.7	255	2	Q5MHQ0_9HIV1	Q5mhq0 human immun
903	34	60.7	238	2	Q6XNY1_9SAUR	Q6xny1 phlesuma ab	976	34	60.7	255	2	Q4TUU2_9HIV1	Q4tuu2 human immun
904	34	60.7	238	2	Q6XNY0_9SAUR	Q6xny0 phlesuma ab	977	34	60.7	256	2	Q4VTU8_9SAUR	Q4vtu8 gymnodactyl
905	34	60.7	238	2	Q6XNS9_9SAUR	Q6xns9 phlesuma gu	978	34	60.7	256	2	Q5RS93_9HIV1	Q5rs93 human immun
906	34	60.7	238	2	Q8AFU0_9HIV1	Q8afu0 human immun	979	34	60.7	256	2	Q5RS99_9HIV1	Q5rs99 human immun
907	34	60.7	238	2			980	34	60.7	257	2		

981 34 60.7 258 2 Q85Q43 trichomya
 982 34 60.7 258 2 Q5RT01 human immun
 983 34 60.7 258 2 Q4UP49 xanthomonas
 984 34 60.7 259 2 Q8P3M5 xanthomonas
 985 34 60.7 259 2 Q8PF17 xanthomonas
 986 34 60.7 259 2 Q8AFU1 human immun
 987 34 60.7 259 2 Q8AF77 human immun
 988 34 60.7 261 2 Q8J877 human immun
 989 34 60.7 261 2 Q5IR78 human immun
 990 34 60.7 262 2 Q9YQ24 chimpanzee
 991 34 60.7 263 2 Q5LVW4 silicibacte
 992 34 60.7 264 2 Q4VTU1 gymnodactyl
 993 34 60.7 264 2 Q5VAC6 human immun
 994 34 60.7 264 2 Q5RQ00 human immun
 995 34 60.7 264 2 Q5RMH8 human immun
 996 34 60.7 266 2 Q34813 isothrix bi
 997 34 60.7 266 2 Q34820 isothrix pa
 998 34 60.7 266 2 Q34421 echinys chr
 999 34 60.7 266 2 Q34812 isothrix bi
 1000 34 60.7 266 2 Q34307 dactylomya

ALIGNMENTS

RESULT 1
 ID Q9RKM5 STROCO PRELIMINARY; PRT; 319 AA.
 AC Q9RKM5
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative MerR family transcriptional regulator.
 GN OrderedLocuNames=SCO4102; ORFNames=SCO17.06c;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939118; CAB56383.1; -; Genomic_DNA.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR000551; HTH_MerR.
 DR Pfam; PF00376; MerR; 1.
 DR PRINTS; PR00040; HTHMER.
 DR SMART; SM00422; HTH_MER; 1.
 DR PROSITE; PS50937; HTH_MER 2; 1.
 KW Complete proteome; DNA-binding.
 SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;

Query Match 78.6%; Score 44; DB 2; Length 319;
 Best Local Similarity 70.0%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EGPTLRQWLA 10
 :|||:|:
 Db 258 DGPRLQWLA 267

RESULT 2
 ID Q7UQ84 RHOBIA PRELIMINARY; PRT; 297 AA.
 AC Q7UQ84;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocuNames=RB6375;
 OS Rhodopirella baltica.
 OC Bacteriia; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.14314443100;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
 RA Schleutner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 DR EMBL; BX294144; CAD74759.1; -; Genomic_DNA.
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR InterPro; IPR003169; GYF.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
 DR PROSITE; PS50829; GYF; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 297 AA; 31805 MW; 475P670F02C78B9B CRC64;

Query Match 76.8%; Score 43; DB 2; Length 297;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 EGPTLRQWLA 10
 :|||:|:
 Db 176 DGPTRQWIS 185

RESULT 3
 ID P90433_SIVCZ PRELIMINARY; PRT; 313 AA.
 AC P90433;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Truncated reverse transcriptase (Fragment).
 GN Name=pol;
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Lentivirus; Primate lentivirus group.
 OX NCBI_TaxID=11723;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Smith J.M., Kraiselburd E.N., Torres J.V.;
 RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 DR EMBL; U83413; AAB41428.1; -; Genomic_DNA.
 DR HSPSP; Q07387; ITCW.
 DR SMR; P90433; 69-167, 173-313.
 DR MEROPS; A02.002; -.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003664; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR InterPro; IPR009007; Pept_Aspartc_cat.
 DR InterPro; IPR001995; Peptidase_A2_cat.

```

DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
DR Aspartyl protease; Hydrolase; Protease; RNA-directed DNA polymerase.
FT NON TER 1
SQ SEQUENCE 313 AA; 34675 MW; 5A0BB016783FC8A6 CRC64;

Query Match 75.0%; Score 42; DB 2; Length 313;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTRLQW 8
Db 184 EGPKLRQW 191

RESULT 4
Q89RH2_BRAJA
ID Q89RH2_BRAJA PRELIMINARY; PRT; 590 AA.
AC Q89RH2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B12800 protein.
GN OrderedLocusNames=bll2800;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBITaxID=375;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=USDA 110;
RC MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamiasawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; BA00040; BAC48065.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00481; PP2C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00333; PP2C_SIG; 1.
DR SMART; SM00332; PP2CC; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN.1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
Transferase.
SQ SEQUENCE 580 AA; 64917 MW; 6AD3A06E6FAE143B CRC64;

Query Match 75.0%; Score 42; DB 2; Length 580;
Best Local Similarity 77.8%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTRLQWL 9
Db 356 EGQTLRQWM 364

RESULT 5
POL_SIVS4

```

```

ID AC POL_SIVS4 STANDARD; PRT; 1019 AA.
DT DT12502;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropesin) (EC 3.4.23.-);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 3.1.26.4) (RT);
DE Integrase (IN)].
GN Names=POL;
OS Simian immunodeficiency virus (isolate F236/emH4) (SIV-SM).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBITaxID=11737;
[1]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RP MEDLINE=89262053; PubMed=2786147; DOI=10.1038/339389a0;
RA Hirsch V.M., Olmstead R.A., Murphey-Corb M., Purcell R.H.,
RA Johnson P.R.;
RT "An African primate lentivirus (SIVsm) closely related to HIV-2.";
RL Nature 339:389-392(1989).
CC -!- FUNCTION: During replicative cycle of retroviruses, the reverse-
transcribed viral DNA is integrated into the host chromosome by
the viral integrase enzyme. RNase H activity is associated with
the reverse transcriptase.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -!- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
diphosphate + DNA(n+1).
CC -!- PM: Cleavage sites that yield the mature proteins remain to be
determined.
CC -!- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
CC -!- SIMILARITY: Contains 1 integrase catalytic domain.
CC -!- SIMILARITY: Contains 1 integrase-type DNA-binding domain.
CC -!- SIMILARITY: Contains 1 integrase-type zinc finger.
CC -!- SIMILARITY: Contains 1 peptidase A2 domain.
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
CC -!- SIMILARITY: Contains 1 RNase H domain.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; X14307; -; NOT_ANNOTATED_CDS; Genomic_DNA.
HSSP; P04584; 1MU2.
SMR; P12502; 69-167, 781-939.
MEROPS; A02.002; -.
HIV; X14307; POL$SMVH4.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Asparts_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_Chumb; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; IEA.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.

```

```

DR PROSITE; PS50878; RT POL; 1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW AIDS; Aspartyl protease; DNA integration; DNA recombination;
KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
KW Nuclease; Nucleotidyltransferase; Polyprotein; Protease;
KW RNA-directed DNA polymerase; Transferrase; Zinc; Zinc-finger.
FT CHAIN 1 167
FT DOMAIN 88 157 Peptidase A2.
FT DOMAIN 211 401 Reverse transcriptase.
FT DOMAIN 600 723 RNase H.
FT DOMAIN 779 930 Integrase catalytic.
FT ZN_FING 729 770 Integrase-type.
FT DNA_BIND 949 996 Integrase-type.
FT ACT_SITE 93 93 By similarity.
SQ SEQUENCE 1019 AA; 115465 MW; 8D3DE0B85FC92B1C CRC64;

Query Match 75.0%; Score 42; DB 1; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTRLQW 8
Db 184 EGPKLQW 191

RESULT 6
P89154_SIVCZ
ID P89154_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC P89154;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Pol protein (Fragment).
GN Name-pol;
OS Chimpanzee immunodeficiency virus (SIV(cp2)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SiVem543;
RX MEDLINE=97151152; PubMed=9995688;
RA Hirsch V., Ager-Johnson D., Campbell B., Goldstein S., Brown C.,
RA Elkins W.K., Montefiori D.C.;
RT "A molecularly cloned, pathogenic, neutralization-resistant simian
RT immunodeficiency virus, SiVem543-3.";
RL J. Virol. 71:1608-1620(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SiVem543;
RA Ourmanov I.K., Deighani H., Kuwata T., Hirsch V.M.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U72748; AAC56559.2; -; Genomic_DNA.
DR SMR; P89154; 69-167, 761-939.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_AspArtc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.

```

```

DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1 1
SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTRLQW 8
Db 184 EGPKLQW 191

RESULT 7
Q7ZBR7_SIVCZ
ID Q7ZBR7_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name-pol;
OS Chimpanzee immunodeficiency virus (SIV(cp2)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1126/JVI.77.11.6405-6418.2003;
RA Deighani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221514; AAC067307.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.

```



```
DR SMR; Q7ZBR7; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00994; ASP_PROTEASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; A886525DFF1BE26F CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred.No.1.le+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTRLQW 8
Db 184 EGFKLQW 191

RESULT 8
Q7ZBR5_SIVCZ
ID Q7ZBR5_SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name-pol;
OS Chimpanzee immunodeficiency virus (SIV(cpx)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
```

```
RX SMR; Q7ZBR7; 69-167, 781-939.
RA "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
RL J. Virol. 77:6405-6418(2003).
CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The LPOG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC EMBL; AY221515; AAO67309.1; -; Genomic_DNA.
DR HSP; P04584; 1MU2.
DR SMR; Q7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase C.
DR InterPro; IPR003308; Integrase N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00994; ASP_PROTEASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 75.0%; Score 42; DB 2; Length 1019;
Best Local Similarity 87.5%; Pred.No.1.le+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 EGPTRLQW 8
Db 184 EGPTRLQW 191

RESULT 9
Q9R7K1_9SPHN PRELIMINARY; PRT; 244 AA.
AC Q9R7K1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrocyte sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3031;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR HMBL; AB015710; BAA32999.1; -; Genomic_DNA.
DR HSP; P02954; IYST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 244 AA; 27110 MW; 1B70EC63A39D95A CRC64;
DR HMBL; AB015710; BAA32999.1; -; Genomic_DNA.
DR HSP; P02954; IYST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 244 AA; 27110 MW; 1B70EC63A39D95A CRC64;

Query Match 73.2%; Score 41; DB 2; Length 244;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTRLQWL 9
Db 27 EGPTRLNPWL 35

RESULT 11
O82991_9SPHN PRELIMINARY; PRT; 245 AA.
AC O82991;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrocyte sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3017;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR HMBL; AB015709; BAA32997.1; -; Genomic_DNA.
DR HSP; P02954; IYST.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. . .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. . .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR Pfam; PF00124; Photo_RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMs; TIGR01157; pufL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1
SQ SEQUENCE 244 AA; 27110 MW; 1B70EC63A39D95A CRC64;

Query Match 73.2%; Score 41; DB 2; Length 244;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTRLQWL 9
Db 22 EGPTRLNPWL 30

RESULT 10
O6272_9SPHN PRELIMINARY; PRT; 245 AA.
AC O6272;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Name=pufL;
OS Erythrocyte litoralis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=39960;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14332;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;

```

```

FT  NON TER      1      1
SQ  SEQUENCE    245 AA;  27229 MW;  6F4A8889BA2192C0 CRC64;

Query Match      73.2%; Score 41; DB 2; Length 245;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  1 EGPTLRQWL 9
Db  23 EGPTLRPWL 31

RESULT 12
O82989_9SPHN
ID  O82989_9SPHN PRELIMINARY; PRT; 249 AA.
AC  O82989;
DT  01-NOV-1998 (TrEMBLrel. 08, Created)
DT  01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Photosynthetic reaction center L subunit (Fragment).
GN  Name=pufL;
OS  Erythrobacter sp.
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC  Sphingomonadaceae; Erythrobacter.
OX  NCBI_TaxID=1042;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=MBIC3019;
RX  MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA  Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA  Hanada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT  "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL  Nature 415:630-633(2002).
DR  EMBL; AB015708; BAA32995.1; -; Genomic_DNA.
DR  HSSP; P02954; 1YST.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. .; IEA.
DR  GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR  GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR  GO; GO:0045156; F:electron transporter, transferring electron. .; IEA.
DR  GO; GO:0000287; F:magnesium ion binding; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR  GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR005871; Photo_RC; 1.
DR  InterPro; IPR000484; Photo_RC.
DR  Pfam; PF00124; Photo_RC; 1.
DR  PRINTS; PR00256; REACTNCENTRE.
DR  TIGRFSMS; TIGR01157; pufL; 1.
DR  PROSITE; PS00244; REACTION_CENTER; 1.
SQ  SEQUENCE 249 AA; 27702 MW; 4D68EDC82B7166AD CRC64;

Query Match      73.2%; Score 41; DB 2; Length 249;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  1 EGPTLRQWL 9
Db  27 EGPTLRPWL 35

RESULT 13
O9XDVO_9SPHN
ID  O9XDVO_9SPHN PRELIMINARY; PRT; 278 AA.
AC  O9XDVO;
DT  01-NOV-1999 (TrEMBLrel. 12, Created)
DT  01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Photosynthetic reaction center L subunit.
GN  Name=pufL;
OS  Erythrobacter sp. MBIC3960.

```

```

OC  Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC  Sphingomonadaceae; Erythrobacter.
OX  NCBI_TaxID=94771;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=MBIC3960;
RX  MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA  Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA  Hanada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT  "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL  Nature 415:630-633(2002).
DR  EMBL; AB027515; BAA78672.1; -; Genomic_DNA.
DR  HSSP; P02954; 1YST.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. .; IEA.
DR  GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR  GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR  GO; GO:0045156; F:electron transporter, transferring electron. .; IEA.
DR  GO; GO:0000287; F:magnesium ion binding; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR  GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR005871; Photo_RC.
DR  InterPro; IPR000484; Photo_RC.
DR  Pfam; PF00124; Photo_RC; 1.
DR  PRINTS; PR00256; REACTNCENTRE.
DR  TIGRFSMS; TIGR01157; pufL; 1.
DR  PROSITE; PS00244; REACTION_CENTER; 1.
SQ  SEQUENCE 278 AA; 30735 MW; 0BE618844B3C54FB CRC64;

Query Match      73.2%; Score 41; DB 2; Length 278;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy  1 EGPTLRQWL 9
Db  56 EGPTLRPWL 64

RESULT 14
Q742B3_MYCPA
ID  Q742B3_MYCPA PRELIMINARY; PRT; 302 AA.
AC  Q742B3;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Galu.
GN  Name=galu; OrderedLocusNames=MAP0924;
OS  Mycobacterium paratuberculosis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC  Mycobacterium avium complex (MAC).
OX  NCBI_TaxID=1770;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=k10;
RA  Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL  Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE017230; AAS03241.1; -; Genomic_DNA.
DR  GO; GO:0016301; F:kinase activity; IEA.
DR  GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0009058; P:biosynthesis; IEA.
DR  InterPro; IPR005835; NTP_transferase.
DR  Pfam; PF00483; NTP_transferase; 1.
KW  Complete proteome.
SQ  SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;

Query Match      73.2%; Score 41; DB 2; Length 302;
Best Local Similarity 87.5%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY 2 GPTLRQWL 9
Db 286 GPDLRQWL 293
||| ||| |||
||| ||| |||

RESULT 15
XERC_CORGL STANDARD; PRT; 308 AA.
AC QNNZ9;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Tyrosine recombinase xerC.
GN Name=xerC; OrderedLocuNames=Cgl12028, cg22224;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bache B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins";
RL J. Biotechnol. 104:5-25(2003).
CC -1- FUNCTION: Site-specific tyrosine recombinase, which acts by
CC catalyzing the cutting and rejoining of the recombining DNA
CC molecules. The xerC-xerD complex is essential to convert dimers of
CC the bacterial chromosome into monomers to permit their segregation
CC at cell division. It also contributes to the segregational
CC stability of plasmids (By similarity).
CC -1- SUBUNIT: Forms a cyclic heterotetrameric complex composed of two
CC molecules of xerC and two molecules of xerD (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the "phage" integrase family. XerC type 1
CC subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC EMBL; BA000036; BAB99421.1; -; Genomic DNA.
CC EMBL; BX927154; CAP20368.1; ALT_INIT; Genomic DNA.
CC HSP; P21891; 1A0P.
CC HAMAP; MF_01808; -; 1.
CC InterPro; IPR004107; Phage_integr_N.
CC InterPro; IPR002104; Phage_integrase.
CC Pfam; PF02899; Phage_integr_N; 1.
CC Pfam; PF00589; Phage_integrase; 1.
CC Cell cycle; Cell division; Chromosome partition; Complete proteome;
CC DNA integration; DNA recombination; DNA-binding.
CC ACT_SITE 163 163 By similarity.
CC ACT_SITE 187 187 By similarity.
CC ACT_SITE 254 254 By similarity.
CC ACT_SITE 257 257 By similarity.
CC ACT_SITE 280 280 By similarity.
CC ACT_SITE 289 289 O-(3'-phospho-DNA)-tyrosine intermediate
CC (By similarity).
FT
```

```
SQ SEQUENCE 308 AA; 34053 MW; 10FA3E71949334FC CRC64;
Query Match 73.2%; Score 41; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRQWL 9
Db 67 PTLRQWL 73
|||||
|||||

RESULT 16
Q617Q3 CAEBR PRELIMINARY; PRT; 728 AA.
AC Q617Q3; 2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14874.
GN Names=CBG14874;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RL The C. briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; CAAC01000068; CAE68902.1; -; Genomic DNA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Ub1 conjugation pathway; Zinc;
KW Zinc-finger.
SQ SEQUENCE 728 AA; 82017 MW; 97016C238A26F96B CRC64;

Query Match 73.2%; Score 41; DB 2; Length 728;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GPTLRQWL 10
Db 614 GPCLRKWL 622
|||||
|||||

RESULT 17
Q95Y82 CAEBL PRELIMINARY; PRT; 754 AA.
AC Q95Y82; 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Y119C1B.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99085613; PubMed=9851916;
RT "The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
```

```
CC -I- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AC006712; AAK39324.1; -; Genomic_DNA.
DR Ensembl; Y119C1B.5; Caenorhabditis elegans.
DR WormBase; WBGene00022471; Y119C1B.5.
DR WormPep; Y119C1B.5; CE27234.
DR GO; GO:0001511; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Complete proteome; Hypothetical protein; Metal-binding;
KW Ubl conjugation pathway; Zinc; Zinc-finger.
SQ SEQUENCE 754 AA; 85324 MW; 41BAA9297FA3BF05 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 754;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GPTLRQWLA 10
   |||||
DB 626 GPCLRWLA 634

RESULT 18
Q73SJ6 MYCPA PRELIMINARY; PRT; 791 AA.
AC Q73SJ6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=MAP4077c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Anonain A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017242; AAS06627.1; -; Genomic_DNA.
DR InterPro; IPR005242; Cons_hypoth374.
DR Pfam; PF03706; UPR0104; 1.
DR TIGRFAMs; TIGR00374; Cons_hypoth374; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 791 AA; 83885 MW; 2E5CB0DE4B2B4F1B CRC64;

Query Match 73.2%; Score 41; DB 2; Length 791;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRWLA 10
   |||||
DB 653 PTLRWLA 660

RESULT 19
Q4KS46_9VIRU PRELIMINARY; PRT; 815 AA.
AC Q4KS46;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tyrosine kinase.
OS Orange-spotted grouper iridovirus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.
OX NCBI_TaxID=322017;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhou S.Y., Lv L., Chen C., Weng S.P., Chan S.M., He J.G.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY894343; AAX82420.1; -; Genomic_DNA.
RW Kinase.
SQ SEQUENCE 815 AA; 93153 MW; 120C2FBS9703F8C4 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 815;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
   :|||||
DB 501 QGPTLAQWI 509

RESULT 20
Q9ST50 MAIZE PRELIMINARY; PRT; 863 AA.
AC Q9ST50;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transposase related protein.
GN Name=trap;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Baumeister P., Koellig J., Werr W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ238507; CAB51950.1; -; Genomic_DNA.
DR InterPro; IPR004332; MuDR.
DR InterPro; IPR006564; Znf_PWZ.
DR InterPro; IPR007527; Znf_SWIM.
DR Pfam; PF03108; MuDR; 1.
DR Pfam; PF04434; SWIM; 1.
DR SMART; SM00575; Znf_PWZ; 1.
DR PROSITE; PS50966; ZF_SWIM; 1.
SQ SEQUENCE 863 AA; 98714 MW; 2B3FDA5C57B172A1 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLRWL 9
   |||||
DB 305 PTLRWL 311

RESULT 21
Q8QUJ6_9VIRU PRELIMINARY; PRT; 941 AA.
AC Q8QUJ6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF114L.
OS Infectious spleen and kidney necrosis virus.
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Megalocytivirus.
OX NCBI_TaxID=180170;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21874810; PubMed=11878882; DOI=10.1006/viro.2001.1208;
RA He J.G., Deng M., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,
RA Chan S.M.;
RT "Complete genome analysis of the mandarin fish infectious spleen and
RT kidney necrosis iridovirus."
RT Virology 291:126-139(2001).
RL
```

```

DR EMBL: AF371960: AAL98838.1; -: Genomic DNA
SQ SEQUENCE 941 AA; 106703 MW; EB663998CF6CE83 CRC64;

Query Match 73.2%; Score 41; DB 2; Length 941;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQWL 9
Db 582 QGPTLAQMI 590

RESULT 22
QSYWJ5_NOCFA PRELIMINARY; PRT; 53 AA.
AC QSYWJ5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=nfa25990;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yanashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL: AP006618; BAB57446.1; -: Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 53 AA; 5806 MW; 47B08B1F9BA4FCE6 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 53;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GPTLRQWLA 10
Db 33 GPQCRQWLA 41

RESULT 23
Q8DHX7_SYNEL PRELIMINARY; PRT; 129 AA.
AC Q8DHX7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T11816 protein.
GN OrderedLocNames=t11816;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BP-1;
RX MEDLINE=22251544; PubMed=12240834;
RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL: BA000039; BAC09368.1; -: Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 129 AA; 14644 MW; EBB44691E7DD1E12 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 129;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GPTLRQWLA 10
Db 57 GPTLRQWRA 65

RESULT 24
Q6FE11_AC1AD PRELIMINARY; PRT; 137 AA.
AC Q6FE11;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=AC1AD0794;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Orsnton L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1 a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
DR EMBL: CR543861; CAG67697.1; -: Genomic_DNA.
DR InterPro; IPR007401; DUF454.
DR Pfam; PF04304; DUF454.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 137 AA; 15749 MW; 70C5D7C68908AFCA CRC64;

Query Match 71.4%; Score 40; DB 2; Length 137;
Best Local Similarity 77.8%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GPTLRQWLA 10
Db 57 GPTLRQWRA 65

RESULT 25
Q54XF9_DICDI PRELIMINARY; PRT; 154 AA.
AC Q54XF9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80215283;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RX Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A., Xu Q.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal B., Kummerfeld S., Maderla N., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Rason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,

```

```

RA Louleaged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shaulesky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RA "the genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAP10100054; EAL67916.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 154 AA; 17998 MW; 2BE3834C15F6FAF6 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 154;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQWL 9
Db 73 QGPTSKQWL 81

RESULT 26
Q8ZGS7 YERPE PRELIMINARY; PRT; 296 AA.
AC Q8ZGS7; Q74WQ0; Q7CH89;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Putative membrane protein (Putative transmembrane protein).
GN Name=xht4; OrderedLocusNames=YPO934, YPO1203, Y2985;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RA "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RA "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans.";

```

```

RL DNA Res. 11:179-197(2004).
DR EMBL: AJ414147; CAC90042.1; -; Genomic_DNA.
DR EMBL: AE013900; AAM86536.1; -; Genomic_DNA.
DR EMBL: AE017130; AAS61189.1; -; Genomic_DNA.
DR PIR: AG0147; AG0147.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016020; C: membrane; IEA.
DR InterPro: IPR000620; DUF6_TM.
DR Pfam: PF00892; DUF6; 2.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 296 AA; 31378 MW; 4594713DCD54CFF CRC64;

Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
Db 66 PTLRQWAA 73

RESULT 27
Q6D06 YERPS PRELIMINARY; PRT; 296 AA.
AC Q6D06;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative drug/metabolite (DME family) efflux pump precursor.
GN OrderedLocusNames=YPTB1243;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RA "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL: BX936398; CAH20483.1; -; Genomic_DNA.
DR GO: GO:0016020; C: membrane; IEA.
DR InterPro: IPR000620; DUF6_TM.
DR Pfam: PF00892; DUF6; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 25 Potential.
SQ SEQUENCE 296 AA; 31407 MW; 4D3E486D32DBAC11 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
Db 66 PTLRQWAA 73

RESULT 28
P95613 RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
GN Name=nodD2;
OS Rhizobium galegae.

```



```

Science 300:1566-1569(2003).
[2]
RN NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Tsitrin T.,
Bera J., Kim M., Jin S., Padrosh D., Vuong H., Overton II L.,
Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
Wortman J., Haas B., Koo H., Zigmann V., Hsiao J., Iobst S.,
de Vazeilles A., White O., Salzberg S., Fraser C.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017067; AAY52546.1; -; Genomic_DNA.
DR EMBL; AC105932; AAX95563.1; -; Genomic_DNA.
DR Gramene; Q8LMK9; -.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein; Polyprotein.
SQ SEQUENCE 760 AA; 82020 MW; C51F91AA2EB32A28 CRC64;

Query Match 71.4%; Score 40; DB 2; Length 760;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWLA 10
Db 663 GPTLRQWMA 671
||||:|:|

RESULT 32
P72684_SVNY3
ID P72684 SVNY3 PRELIMINARY; PRT; 131 AA.
AC P72684;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Slr0740 protein.
GN OrderedLocusNames=slr0740;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN NUCLEOTIDE SEQUENCE.
RA Tabata S.;
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyaajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996)
DR EMBL; BA000022; BAA16691.1; -; Genomic_DNA.
DR PIR; S74539; S74539.
KW Complete proteome.
SQ SEQUENCE 131 AA; 14495 MW; DB088997ACF2CB70 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 131;
Best Local Similarity 87.5%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLRQWLA 10
Db 29 PQLRQWLA 36
|||||||

```

```

RESULT 33
Q42PV7_PSESY PRELIMINARY; PRT; 133 AA.
AC Q42PV7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Psyr 3784;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=B728a;
RC DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
Nolan M., Goldsman E., Thiel J., Malfatti S., Lapidus A., Detter J.C.,
Land M., Richardson P.M., Kyrpides N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
pv. syringae B728a and pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=B728a;
RC DOE Joint Genome Institute;
RA Loper J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=B728a;
RC DOE Joint Genome Institute;
RA Feil H., Feil W.S., Lindow S.E.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000075; AAY38815.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 133 AA; 14404 MW; 4176926884998F43 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 133;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
Db 28 GPTYREWL 35
|||||:|:|

RESULT 34
Q88615_PSESM PRELIMINARY; PRT; 134 AA.
AC Q88615;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PSPT01594;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
Grimm M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
Nelson W.C., Davidssen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen

```

```
RT Pseudomonas syringae pv. tomato DC3000. ";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016853; AA055114.1; -; Genomic_DNA.
DR TIGR; PSPT01594; -.
SQ SEQUENCE 134 AA; 14322 MW; 81A064617211C28E CRC64;

Query Match 69.6%; Score 39; DB 2; Length 134;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
Db 28 GPTVREWL 35

RESULT 35
Q9WI01_9HIV2
ID Q9WI01_9HIV2 PRELIMINARY; PRT; 154 AA.
AC Q9WI01;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protease (Fragment).
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP MEDLINE=20211659; PubMed=10747109;
RA Rodas B., Holguin A., Soriano V., Dourana M., Mansinho K., Antunes F.,
RA Gonzalez-Lahoz J.;
RT "Emergence of drug resistance mutations in human immunodeficiency
RT virus type 2-infected subjects undergoing antiretroviral therapy.";
RL J. Clin. Microbiol. 38:1370-1374(2000).
DR EMBL; AF139053; AAD37335.1; -; Genomic_DNA.
DR HSP; P04584; 1IDA.
DR SMK; Q9WI01; 1-98, 101-154.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001955; Peptidase A2_cat.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00077; RVP; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR KW Aspartyl1 protease; Hydrolase; Protease.
FT NON_TER 1
FT NON_TER 154
SQ SEQUENCE 154 AA; 16821 MW; C47EB4C660405791 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 154;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 115 DGPKLRQW 122

RESULT 36
Q6MNC9_BDEBA
ID Q6MNC9_BDEBA PRELIMINARY; PRT; 164 AA.
AC Q6MNC9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedLocusNames=Ed1329;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
```

```
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: Life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842649; CAE79223.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1
FT SIGNAL 7
SQ SEQUENCE 164 AA; 17624 MW; D9B7CCFDDA0246A2 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 164;
Best Local Similarity 60.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EGBTLRQWLA 10
Db 60 QGPALQOWTA 69

RESULT 37
Q9V492_DROME
ID Q9V492_DROME PRELIMINARY; PRT; 168 AA.
AC Q9V492;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG11077-PA (R55125p).
GN ORFNames=CG11077;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherzer S.E., Li P.W., Hoskins R.A., Galle R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartie N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
```

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celisner S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park P., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor S.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celisner S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley *Drosophila* Genome Project;
RA Celisner S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirkas R., Smith E.,
RA Yu C., Rubin G.;
RT "*Drosophila melanogaster* release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celisner S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003846; AAF59393.1; -; Genomic DNA.
DR EMBL; AY071482; AAL49104.1; -; mRNA.
DR Ensembl; CG11077; *Drosophila melanogaster*.
DR FlyBase; FSN003930; CG11077.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR007248; Mpv17_PMP22.
DR Pfam; PF04117; Mpv17_PMP22; 1.
SQ SEQUENCE 168 AA; 19521 MW; 48E216A954E43D39 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 168;

Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 GPTLRQW 8
DB 54 GPTLRW 60
RESULT 38
Q87115 SIVCZ PRELIMINARY; PRT; 217 AA.
AC Q87115
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SiVagmsABD37;
RX MEDLINE=94298785; PubMed=8026477;
RA Jin M.J., Hui H., Robertson D.L., Muller M.C., Barre-Sinoussi F.,
RA Hirsch V.M., Allan J.S., Shaw G.M., Sharp P.M., Hahn B.H.;
RT "Mosaic genome structure of simian immunodeficiency virus from west
RT African green monkeys".
RL EMBO J. 13:2935-2947(1994).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U04018; AAA21512.1; -; Genomic DNA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR009007; Peptidase_A2_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS0175; ASP_PROT_RETROV; 1.
DR PROSITE; PS0878; RT_POL; 1.
FT NON_TER 1
FT NON_TER 217
SQ SEQUENCE 217 AA; 24503 MW; C1162B4BE18204B8 CRC64;
Query Match 69.6%; Score 39; DB 2; Length 217;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 EGPTLRQW 8
DB 86 DGPTLRQW 93
RESULT 39
Q8RJN8 RAT PRELIMINARY; PRT; 218 AA.
AC Q8RJN8
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical LOC287938.
GN Name=LOC287938;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Testis; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX MEDLINE=22388257; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan T., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uadin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG NIH MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC086565; AA086555.1; -; mRNA.
DR InterPro; IPR006879; YdJC.
DR Pfam; PF04794; YdJC; 1.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 23436 MW; 7BCCAF3DD2B6ABCE CRC64;

Query Match 69.6%; Score 39; DB 2; Length 218;
Best Local Similarity 87.5%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 PTLRWLWLA 10
DB 188 PTLRAWLA 195

RESULT 40
Q7ZKL1_9HIV1
ID Q7ZKL1_9HIV1 PRELIMINARY; PRT; 230 AA.
AC Q7ZKL1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 1.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11676;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22866097; PubMed=14506786; DOI=10.1089/08992220322280946;
RA Toni T.D., Recordon-Pinson P., Minga A., Ekouevi D., Bonard D.,
RA Bequet L., Huet C., Chenal H., Rouet F., Dabis F., Lafon M.-E.,
RA Salamon R., Masquelier B., Fleury H.J.;
RT "Presence of key drug resistance mutations in isolates from untreated
RT patients of Abidjan, Cote d'Ivoire : ANRS 1257 Study";
RL AIDS Res. Hum. Retroviruses 19:713-717(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fleury H., Toni T., Recordon-Pinson P., Minga A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY207863; AA053469.1; -; Genomic_DNA.
DR SMR; Q7ZKL1; 1-230.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.

DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR000477; RVTS.
DR Pfam; PF00078; RVT 1; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 230 230
SQ SEQUENCE 230 AA; 26696 MW; 0FBDBA8B46D5952C CRC64;

Query Match 69.6%; Score 39; DB 2; Length 230;
Best Local Similarity 82.5%; Pred. No. 85;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGPTLRQW 8
DB 3 DGPTVQW 10

RESULT 41
O82987_9SPHN
ID O82987_9SPHN PRELIMINARY; PRT; 245 AA.
AC O82987;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Photosynthetic reaction center L subunit (Fragment).
GN Names=puL;
OS Erythrobacter sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=1042;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBIC3363;
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., Delong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs";
RL Nature 415:630-633 (2002).
DR EMBL; AB015707; BAA32993.1; -; Genomic_DNA.
DR HSSP; P02954; IPSS.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0030077; C:light-harvesting complex (sensu Proteobacte. .; IEA.
DR GO; GO:0030090; C:reaction center (sensu Proteobacteria); IEA.
DR GO; GO:0042314; F:bacteriochlorophyll binding; IEA.
DR GO; GO:0045156; F:electron transporter, transferring electron. .; IEA.
DR GO; GO:0000287; F:magnesium ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0019684; P:photosynthesis, light reaction; IEA.
DR GO; GO:0018298; P:protein-chromophore linkage; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005871; Photo_L.
DR InterPro; IPR000484; Photo_RC.
DR Pfam; PF00124; Photo RC; 1.
DR PRINTS; PR00256; REACTNCENTRE.
DR TIGRFAMS; TIGR01157; puFL; 1.
DR PROSITE; PS00244; REACTION_CENTER; 1.
FT NON_TER 1 1
FT NON_TER 245 245
SQ SEQUENCE 245 AA; 27411 MW; 06C24DD9610CF92D CRC64;

Query Match 69.6%; Score 39; DB 2; Length 245;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGPTLRQWL 9
DB 22 EGPTMNPWL 30

RESULT 42
Q5YPT6_NOCFA
ID Q5YPT6_NOCFA PRELIMINARY; PRT; 298 AA.
AC Q5YPT6;

DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Putative UDP-glucose-1-phosphate uridylyltransferase.
 GN Name=galu; OrderedLocusNames=nfa49530;
 OS Nocardia farcinica.
 NC Nocardia farcinica.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Nocardiae; Nocardia.
 OX NCBITaxID=37329;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IFM 10152;
 RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
 RA Iehikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
 RL EMBL; AP006618; BAD59805.1; -; Genomic DNA.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0009038; P:biosynthesis; IEA.
 DR InterPro; IPR005835; NTP transferase.
 DR Pfam; PF00483; NTP transferase; 1.
 KW Complete proteome; Nucleotidyltransferase; Transferase.
 SQ SEQUENCE 298 AA; 31931 MW; F29B1A16B26ABA0D CRC64;

Query Match 69.6%; Score 39; DB 2; Length 298;
 Best Local Similarity 75.0%; Pred. NO. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
 Db 275 GPALREWL 282

RESULT 43
 Q9D8Z7_MOUSE
 ID Q9D8Z7_MOUSE PRELIMINARY; PRT; 310 AA.
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched
 DE library, clone:1810015A11 product:ubiquitin-conjugating enzyme E2L 3,
 DE full insert sequence.
 GN Name=1810015A11Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBITaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito K., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RA Konno H., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Itoh K., Togawa Y., Iwata M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Itoh K., Togawa Y., Iwata M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Pancreas;
 RA Adachi J., Aizawa K., Akahira S., Fukuda S., Fukunishi Y., Furuno M.,
 RA Arakawa T., Bono H., Carninci P., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai T.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK007503; BAB25075.1; -; mRNA.
 DR Ensembl; ENSMUSG00000041774; Mus musculus.
 DR MGI; MGI:1916351; 1810015A11Rik.
 DR InterPro; IPR006879; YdJC.
 DR Pfam; PF04794; YdJC; 1.
 SQ SEQUENCE 310 AA; 33094 MW; 9D3302745CC24048 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 310;
 Best Local Similarity 87.5%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PTLROWLA 10
 Db 280 PTLRAWLA 287

RESULT 44

```

Q7SKI9_9HIV2
ID Q7SKI9_9HIV2 PRELIMINARY; PRT; 310 AA.
AC Q7SKI9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN NamePol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/089922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gleslen M., Tuveson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
RT antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324219; AAP85282.1; -; Genomic_DNA.
DR HSP; P04584; 1IDA.
DR SMR; Q7SKI9; 1-99, 105-310.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Rept_ASP_AS.
DR InterPro; IPR009007; Rept_ASPart_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR00477; RVtse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS01075; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
DR Aspartyl protease; Hydrolase; Polypeptide; Protease.
FT NON_TER 1
FT TER 310
SQ SEQUENCE 310 AA; 35077 MW; 33816141D95E1FD1 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 310;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 116 DGPKLQW 123
|||:|||||
|:::|

RESULT 45
Q4PIZ3_USTWA
ID Q4PIZ3_USTWA PRELIMINARY; PRT; 338 AA.
AC Q4PIZ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM05870.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelheil A., Adekoya E.,
RA Alt-zahra M., Allen T., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,

```

```

RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hegos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysselellis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Maraballa R., Maru K., Matthews C., Maucell E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Purcell S., Rachupka T., Ramasamy U., Rodriguez J., Rogers J., Rogov P.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Staiker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tensing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Teamlia T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACP0100215; EAK86815.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 37245 MW; 838DC845D4ACC448 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 338;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GPTLRQWL 9
Db 294 GPSLRKWL 301
|||:|||||
|:::|

RESULT 46
Q8UN04_SIVCZ
ID Q8UN04_SIVCZ PRELIMINARY; PRT; 340 AA.
AC Q8UN04;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Namespol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN NUCLEOTIDE SEQUENCE.
RP Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF458219; AAL59619.1; -; Genomic_DNA.
DR HSP; P04584; 1MU2.
DR SMR; Q8UN04; 3-340.

```

```
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: 0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro: IPR010661; RVT_thumb.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT_1; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 340
FT NON_TER 340
SQ SEQUENCE 340 AA; 39546 MW; F9F3BFD3F4005252 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 17 DGPKLQW 24

RESULT 47
Q8UN03_SIVCZ
ID Q8UN03_SIVCZ PRELIMINARY; PRT; 340 AA.
AC Q8UN03;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Reverse transcriptase (Fragment).
GN Name=pol;
OS Chimpazee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Diamond T.L., Lee K.Y., Kimata J.T., Kim B.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AF458220; AAL59620.1; -; Genomic_DNA.
DR HSP; P04584; 1MU2.
DR SMR; Q8UN03; 3-340.
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: 0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro: IPR010661; RVT_thumb.
DR InterPro: IPR000477; RVTse.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW RNA-directed DNA polymerase.
FT NON_TER 340
FT NON_TER 340
SQ SEQUENCE 340 AA; 39548 MW; 7777BFD2A057EA6B CRC64;

Query Match 69.6%; Score 39; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 17 DGPKLQW 24

RESULT 48
Q7SKK9_9HIV2
ID Q7SKK9_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK9;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
FT NON_TER 340
FT NON_TER 340
SQ SEQUENCE 340 AA; 39548 MW; 7777BFD2A057EA6B CRC64;

Query Match 69.6%; Score 39; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 17 DGPKLQW 24

RESULT 49
Q7SKK8_9HIV2
ID Q7SKK8_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK8;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/0892220332230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveesson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324200; AAP85263.1; -; Genomic_DNA.
DR HSP; P04584; 1IDA.
DR SMR; Q7SKK8; 1-99, 102-349.
DR GO: 0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: 0008233; F:peptidase activity; IEA.
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: 0006508; P:proteolysis and peptidolysis; IEA.
DR GO: 0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR001969; Pept_Asp_AS.
DR InterPro: IPR009007; Pept_Aspartc_cat.
DR InterPro: IPR001995; Peptidase_A2_cat.
DR InterPro: IPR000477; RVTse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Polyprotein; Protease.
FT NON_TER 349
FT NON_TER 349
SQ SEQUENCE 349 AA; 39974 MW; 2B66AE503EFC35B8 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTLRQW 8
Db 116 DGPKLQW 123

RESULT 49
Q7SKK8_9HIV2
ID Q7SKK8_9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK8;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; PubMed=12908931; DOI=10.1089/0892220332230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gisslen M., Tuveesson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324200; AAP85263.1; -; Genomic_DNA.
DR HSP; P04584; 1IDA.
DR SMR; Q7SKK8; 1-99, 102-349.
DR GO: 0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO: 0008233; F:peptidase activity; IEA.
```

```
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Polypeptide; Protease.
FT NON_TER 1
FT NON_TER 349
SQ SEQUENCE 349 AA; 39855 MW; 82DAEF2DC13E6DED CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTRLQW 8
Db 116 DGPKLQW 123

RESULT 50
Q7SKK6.9HIV2
ID Q7SKK6 9HIV2 PRELIMINARY; PRT; 349 AA.
AC Q7SKK6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol polyprotein (Fragment).
GN Name=pol;
OS Human immunodeficiency virus 2.
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11709;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22790437; Pubmed=12908931; DOI=10.1089/088922203322230905;
RA Brandin E., Lindborg L., Gyllenstein K., Brostrom C., Hagberg L.,
RA Gasslen M., Tuveson B., Blaxhult A., Albert J.;
RT "pol gene sequence variation in Swedish HIV-2 patients failing
RT antiretroviral therapy.";
RL AIDS Res. Hum. Retroviruses 19:543-550(2003).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY324202; AAP85265.1; -; Genomic_DNA.
DR HSP; P04584; IIDA.
DR SMR; Q7SKK6; 1-99, 102-349.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00878; RT_POL; 1.
KW Aspartyl protease; Hydrolase; Polypeptide; Protease.
FT NON_TER 1
FT NON_TER 349
SQ SEQUENCE 349 AA; 39830 MW; 593B83A7F9FFA8D0 CRC64;

Query Match 69.6%; Score 39; DB 2; Length 349;
```

```
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EGPTRLQW 8
Db 116 DGPKLQW 123

Search completed: May 12, 2006, 10:51:45
Job time : 87.7692 secs
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 101.709 Seconds

(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-67

Perfect score: 81

Sequence: 1 CELVGPLMSWLTC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

A_Geneseq_21.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	14	2 AAW09483	Thrombopo
2	81	100.0	14	2 AAW36634	Thrombopo
3	81	100.0	14	2 AAW36767	Thrombopo
4	81	100.0	14	3 AAB17011	TPO-mimet
5	81	100.0	14	4 AAU25853	Human thr
6	81	100.0	14	5 ABB72897	TPO mimet
7	81	100.0	14	7 ADJ73048	TPO mimet
8	81	100.0	14	8 ADJ52683	CHI delet
9	81	100.0	14	8 ADJ51644	CHI delet
10	53	65.4	18	7 ADN59675	Thrombopo
11	53	65.4	25	7 ADN59730	Thrombopo
12	52	64.2	14	3 AAB17010	TPO-mimet
13	52	64.2	14	5 ABB72896	TPO mimet
14	52	64.2	14	7 ADJ73047	TPO mimet
15	52	64.2	14	8 ADJ52682	CHI delet
16	52	64.2	14	8 ADJ51643	CHI delet
17	49	60.5	129	5 ABB09517	Human RNA
18	49	60.5	129	4 AAM29099	Peptide #
19	49	60.5	199	4 AAU27672	Human pro
20	49	60.5	199	4 AAG77853	Human rib
21	49	60.5	199	4 AAU29085	Human PRO
22	49	60.5	199	4 AAG89149	Human sec
23	49	60.5	199	5 ABP43532	Human sec
24	49	60.5	199	5 AAU12076	Human rib

25	49	60.5	199	6	ABU58693	Human PRO
26	49	60.5	199	6	ABU88241	Novel hum
27	49	60.5	199	6	ABU84556	Human sec
28	49	60.5	199	6	ABR66430	Human sec
29	49	60.5	199	6	ABR65820	Human sec
30	49	60.5	199	6	ABU99760	Human PRO
31	49	60.5	199	6	ABU82999	Human PRO
32	49	60.5	199	6	ABU90120	Novel hum
33	49	60.5	199	6	ABR68369	Human sec
34	49	60.5	199	6	ABR64422	Novel hum
35	49	60.5	199	6	ABU92853	Human sec
36	49	60.5	199	6	ABO08930	Human sec
37	49	60.5	199	6	ABO02982	Human sec
38	49	60.5	199	6	ABR75136	Human sec
39	49	60.5	199	6	ABR94898	Human sec
40	49	60.5	199	6	ABU85871	Human PRO
41	49	60.5	199	6	ABU99031	Novel hum
42	49	60.5	199	6	ABU98246	Novel hum
43	49	60.5	199	6	ABU91952	Novel hum
44	49	60.5	199	6	ABU89645	Human PRO
45	49	60.5	199	6	ABU86486	Human sec
46	49	60.5	199	6	ABU67699	Human sec
47	49	60.5	199	6	ABU80727	Human PRO
48	49	60.5	199	6	ABR99645	Human sec
49	49	60.5	199	6	ABR99035	Human sec
50	49	60.5	199	6	ABO16558	Human sec
51	49	60.5	199	6	ABR92458	Human sec
52	49	60.5	199	6	ABO19099	Human sec
53	49	60.5	199	6	ABR78520	Human sec
54	49	60.5	199	6	ABU85256	Novel hum
55	49	60.5	199	6	ABO00395	Novel hum
56	49	60.5	199	6	ABO11727	Human sec
57	49	60.5	199	6	ABO02372	Human sec
58	49	60.5	199	6	ABU88946	Novel hum
59	49	60.5	199	6	ABU83641	Human sec
60	49	60.5	199	6	ABO08442	Novel hum
61	49	60.5	199	6	ABR59478	Human sec
62	49	60.5	199	6	ABO09540	Human sec
63	49	60.5	199	6	ABO19404	Novel hum
64	49	60.5	199	6	ABO11422	Human sec
65	49	60.5	199	6	ABR67040	Human sec
66	49	60.5	199	6	ABO16253	Human sec
67	49	60.5	199	6	ABO13959	Human sec
68	49	60.5	199	6	ABU65862	Human sec
69	49	60.5	199	6	ABO07710	Human PRO
70	49	60.5	199	6	ABO03897	Human sec
71	49	60.5	199	6	ABR67345	Human sec
72	49	60.5	199	6	ABO15948	Human sec
73	49	60.5	199	6	ABU56229	Human sec
74	49	60.5	199	6	ABU65557	Human PRO
75	49	60.5	199	6	ABU95502	Novel hum
76	49	60.5	199	6	ABU71405	Human PRO
77	49	60.5	199	6	ABO08015	Human PRO
78	49	60.5	199	6	ABR70256	Human sec
79	49	60.5	199	6	ABR69589	Human sec
80	49	60.5	199	6	ABO01730	Human PRO
81	49	60.5	199	6	ABU81532	Human PRO
82	49	60.5	199	6	ABR60329	Human PRO
83	49	60.5	199	6	ABR68064	Human sec
84	49	60.5	199	6	ABR65452	Human sec
85	49	60.5	199	6	ABR68674	Human sec
86	49	60.5	199	6	ABR72086	Human sec
87	49	60.5	199	6	ABU85566	Human PRO
88	49	60.5	199	6	ABU89256	Human sec
89	49	60.5	199	6	ABU83336	Human sec
90	49	60.5	199	6	ABU83336	Novel hum
91	49	60.5	199	6	ABU95192	Novel hum
92	49	60.5	199	6	ABU90740	Novel hum
93	49	60.5	199	6	ABU84251	Human sec
94	49	60.5	199	6	ABU93902	Novel hum
95	49	60.5	199	6	ABR65147	Human sec
96	49	60.5	199	6	ABR68979	Human sec
97	49	60.5	199	6	ABO06795	Human sec
				6	ABR99340	Human sec

98	49	60.5	199	6	ABU57224	Abu57224 Human PRO	171	49	60.5	199	6	ABM35307	Human sec
99	49	60.5	199	6	ABU86176	Abu86176 Novel hum	172	49	60.5	199	6	ABM26544	Human sec
100	49	60.5	199	6	ABU82463	Abu82463 Novel hum	173	49	60.5	199	6	ABO48326	Human sec
101	49	60.5	199	6	ABU87474	Abu87474 Human PRO	174	49	60.5	199	6	ABR93068	Human sec
102	49	60.5	199	6	ABU83946	Abu83946 Human sec	175	49	60.5	199	6	ABO24829	Human sec
103	49	60.5	199	6	ABO08320	Abu08320 Human PRO	176	49	60.5	199	6	ABM11840	Human sec
104	49	60.5	199	6	ABU82031	Abu82031 Novel hum	177	49	60.5	199	6	ABM02941	Human sec
105	49	60.5	199	6	ABU66195	Abu66195 Novel hum	178	49	60.5	199	6	ABM16237	Human sec
106	49	60.5	199	6	ABR60024	Abu60024 Human sec	179	49	60.5	199	6	ABO27798	Human sec
107	49	60.5	199	6	ABU94212	Abu94212 Novel hum	180	49	60.5	199	6	ABM29289	Human sec
108	49	60.5	199	6	ABO00085	Abu000085 Novel hum	181	49	60.5	199	6	ABM07265	Human sec
109	49	60.5	199	6	ABR66735	Abu66735 Human sec	182	49	60.5	199	6	ABM21359	Human sec
110	49	60.5	199	6	ABR91153	Abu91153 Human sec	183	49	60.5	199	6	ABM09705	Human sec
111	49	60.5	199	6	ABU94580	Abu94580 Human PRO	184	49	60.5	199	6	ABO41575	Human sec
112	49	60.5	199	6	ABU79462	Abu79462 Human PRO	185	49	60.5	199	6	ABO36390	Human PRO
113	49	60.5	199	6	ABU86791	Abu86791 Human sec	186	49	60.5	199	6	ABO43919	Human PRO
114	49	60.5	199	6	ABU87096	Abu87096 Novel hum	187	49	60.5	199	6	ABM76619	Human sec
115	49	60.5	199	6	ABU94885	Abu94885 Human PRO	188	49	60.5	199	6	ABM76315	Human sec
116	49	60.5	199	6	ABO04812	Abu04812 Human PRO	189	49	60.5	199	6	ABM25934	Human sec
117	49	60.5	199	6	ABR70561	Abu70561 Human sec	190	49	60.5	199	6	ABM26239	Human sec
118	49	60.5	199	6	ABU98726	Abu98726 Human PRO	191	49	60.5	199	6	ABO03592	Human sec
119	49	60.5	199	6	ABR66125	Abu66125 Human sec	192	49	60.5	199	6	ABO02677	Human sec
120	49	60.5	199	6	ABR64842	Abu64842 Human sec	193	49	60.5	199	6	ABR90848	Human sec
121	49	60.5	199	6	ABU9767	Abu9767 Human PRO	194	49	60.5	199	6	ABR73916	Human sec
122	49	60.5	199	6	ABU93158	Abu93158 Human sec	195	49	60.5	199	6	ABO17168	Human sec
123	49	60.5	199	6	ABU96117	Abu96117 Human PRO	196	49	60.5	199	6	ABR94593	Human sec
124	49	60.5	199	6	ABU91337	Abu91337 Novel hum	197	49	60.5	199	6	ABR76100	Human sec
125	49	60.5	199	6	ABU90430	Abu90430 Novel hum	198	49	60.5	199	6	ABR71476	Human sec
126	49	60.5	199	6	ABO09845	Abu09845 Human sec	199	49	60.5	199	6	ABR93373	Human sec
127	49	60.5	199	6	ABO11117	Abu11117 Human sec	200	49	60.5	199	6	ABR93678	Human sec
128	49	60.5	199	6	ABR71171	Abu71171 Human sec	201	49	60.5	199	6	ABR88103	Human sec
129	49	60.5	199	6	ABU87779	Abu87779 Human PRO	202	49	60.5	199	6	ABO28103	Human sec
130	49	60.5	199	6	ABU91647	Abu91647 Human PRO	203	49	60.5	199	6	ABO30238	Human sec
131	49	60.5	199	6	ABU84861	Abu84861 Human sec	204	49	60.5	199	6	ABO33447	Human PRO
132	49	60.5	199	6	ABR69951	Abu69951 Human sec	205	49	60.5	199	6	ABM05135	Human sec
133	49	60.5	199	6	ABU80328	Abu80328 Human PRO	206	49	60.5	199	6	ABM09095	Human sec
134	49	60.5	199	6	ABU93597	Abu93597 Human PRO	207	49	60.5	199	6	ABO36695	Human sec
135	49	60.5	199	6	ABO10150	Abu10150 Human sec	208	49	60.5	199	6	ABO35780	Human PRO
136	49	60.5	199	6	ABO09235	Abu09235 Human sec	209	49	60.5	199	6	ABO39745	Human sec
137	49	60.5	199	6	ABU10803	Abu10803 Human sec	210	49	60.5	199	6	ABM10620	Human sec
138	49	60.5	199	6	ABU95812	Abu95812 Human PRO	211	49	60.5	199	6	ABM12145	Human sec
139	49	60.5	199	6	ABU97021	Abu97021 Novel hum	212	49	60.5	199	6	ABO52291	Human PRO
140	49	60.5	199	6	ABR70866	Abu70866 Human sec	213	49	60.5	199	6	ABO52596	Human PRO
141	49	60.5	199	6	ABO05217	Abu05217 Novel hum	214	49	60.5	199	6	ABO23914	Human sec
142	49	60.5	199	6	ABO08625	Abu08625 Human sec	215	49	60.5	199	6	ABR97400	Human sec
143	49	60.5	199	6	ABO05832	Abu05832 Human sec	216	49	60.5	199	6	ABU63384	Human rib
144	49	60.5	199	6	ABR74221	Abu74221 Human sec	217	49	60.5	199	6	ABR87188	Human sec
145	49	60.5	199	6	ABR95813	Abu95813 Human sec	218	49	60.5	199	6	ABM11230	Human sec
146	49	60.5	199	6	ABR81110	Abu81110 Human sec	219	49	60.5	199	6	ABM28374	Human sec
147	49	60.5	199	6	ABR81415	Abu81415 Human sec	220	49	60.5	199	6	ABO32373	Human sec
148	49	60.5	199	6	ABM01111	Abu01111 Human sec	221	49	60.5	199	6	ABM15500	Human sec
149	49	60.5	199	6	ABR88713	Abu88713 Human sec	222	49	60.5	199	6	ABM06655	Human sec
150	49	60.5	199	6	ABM77534	Abu77534 Human sec	223	49	60.5	199	6	ABM04466	Human sec
151	49	60.5	199	6	ABO29018	Abu29018 Human sec	224	49	60.5	199	6	ABM22579	Human sec
152	49	60.5	199	6	ABO31763	Abu31763 Human sec	225	49	60.5	199	6	ABM07875	Human sec
153	49	60.5	199	6	ABM08180	Abu08180 Human sec	226	49	60.5	199	6	ABO40965	Human sec
154	49	60.5	199	6	ABO40660	Abu40660 Human sec	227	49	60.5	199	6	ABO40965	Human sec
155	49	60.5	199	6	ABO36085	Abu36085 Human PRO	228	49	60.5	199	6	ABM33375	Human sec
156	49	60.5	199	6	ABO44224	Abu44224 Human PRO	229	49	60.5	199	6	ABO52901	Human PRO
157	49	60.5	199	6	ADA78340	Abu78340 Human sec	230	49	60.5	199	6	ABO50461	Human sec
158	49	60.5	199	6	ABM25019	Abu25019 Human sec	231	49	60.5	199	6	ABU99455	Human sec
159	49	60.5	199	6	ABO03287	Abu03287 Human sec	232	49	60.5	199	6	ABO04507	Human sec
160	49	60.5	199	6	ABR90543	Abu90543 Human sec	233	49	60.5	199	6	ABM18677	Human sec
161	49	60.5	199	6	ABM17457	Abu17457 Human sec	234	49	60.5	199	6	ABR97705	Human sec
162	49	60.5	199	6	ABR95203	Abu95203 Human sec	235	49	60.5	199	6	ABR80805	Human sec
163	49	60.5	199	6	ABR95508	Abu95508 Human sec	236	49	60.5	199	6	ABM01416	Human sec
164	49	60.5	199	6	ABO21746	Abu21746 Human sec	237	49	60.5	199	6	ABR89018	Human sec
165	49	60.5	199	6	ABR98010	Abu98010 Human sec	238	49	60.5	199	6	ABM13670	Human sec
166	49	60.5	199	6	ABR87798	Abu87798 Human sec	239	49	60.5	199	6	ABM21054	Human sec
167	49	60.5	199	6	ABM77839	Abu77839 Human sec	240	49	60.5	199	6	ABO42185	Human sec
168	49	60.5	199	6	ABM28069	Abu28069 Human sec	241	49	60.5	199	6	ABO42795	Human sec
169	49	60.5	199	6	ABM06350	Abu06350 Human sec	242	49	60.5	199	6	ABM10315	Human sec
170	49	60.5	199	6	ABM03856	Abu03856 Human sec	243	49	60.5	199	6	ABO38830	Human sec

244	49	60.5	199	6	ABM33070	Human sec	317	49	60.5	199	6	ABM75705	Human sec
245	49	60.5	199	6	ABM22884	Human sec	318	49	60.5	199	6	ABM25629	Human sec
246	49	60.5	199	6	ABM75095	Human sec	319	49	60.5	199	6	ABM20139	Human sec
247	49	60.5	199	6	ADA80132	Human sec	320	49	60.5	199	6	ABO47045	Human PRO
248	49	60.5	199	6	ABR96485	Human sec	321	49	60.5	199	6	ABO47350	Human PRO
249	49	60.5	199	6	ABM02636	Human sec	322	49	60.5	199	6	ADA83657	Human sec
250	49	60.5	199	6	ABR86578	Human sec	323	49	60.5	199	6	ABR71781	Human sec
251	49	60.5	199	6	ABR86883	Human sec	324	49	60.5	199	6	ABR72391	Human sec
252	49	60.5	199	6	ABM16847	Human sec	325	49	60.5	199	6	ABR98730	Human sec
253	49	60.5	199	6	ABM29899	Human sec	326	49	60.5	199	6	ABO07100	Human sec
254	49	60.5	199	6	ABO29323	Human sec	327	49	60.5	199	6	ABR85053	Human sec
255	49	60.5	199	6	ABM24104	Human sec	328	49	60.5	199	6	ABR73611	Human sec
256	49	60.5	199	6	ABM23494	Human sec	329	49	60.5	199	6	ABR76705	Human sec
257	49	60.5	199	6	ABM22274	Human sec	330	49	60.5	199	6	ABR73306	Human sec
258	49	60.5	199	6	ABO37915	Human sec	331	49	60.5	199	6	ABM18372	Human sec
259	49	60.5	199	6	ABM28679	Human sec	332	49	60.5	199	6	ABO20831	Human sec
260	49	60.5	199	6	ABM28984	Human sec	333	49	60.5	199	6	ABO25574	Human PRO
261	49	60.5	199	6	ABM66628	Human sec	334	49	60.5	199	6	ABO25879	Human PRO
262	49	60.5	199	6	ABM76010	Human sec	335	49	60.5	199	6	ABR94288	Human sec
263	49	60.5	199	6	ABM34290	Human sec	336	49	60.5	199	6	ABR80195	Human sec
264	49	60.5	199	6	ABM34595	Human sec	337	49	60.5	199	6	ABM11535	Human sec
265	49	60.5	199	6	ABO20526	Human sec	338	49	60.5	199	6	ABO33142	Human PRO
266	49	60.5	199	6	ABO21441	Human sec	339	49	60.5	199	6	ABO30848	Human sec
267	49	60.5	199	6	ABO22356	Human sec	340	49	60.5	199	6	ABO31153	Human sec
268	49	60.5	199	6	ABR96790	Human sec	341	49	60.5	199	6	ABM27459	Human sec
269	49	60.5	199	6	ABR85968	Human sec	342	49	60.5	199	6	ABM30204	Human sec
270	49	60.5	199	6	ABR99950	Human sec	343	49	60.5	199	6	ABM05740	Human sec
271	49	60.5	199	6	ABM00501	Human sec	344	49	60.5	199	6	ABM15805	Human sec
272	49	60.5	199	6	ABO20806	Human sec	345	49	60.5	199	6	ABM08790	Human sec
273	49	60.5	199	6	ABO29933	Human sec	346	49	60.5	199	6	ABO42490	Human sec
274	49	60.5	199	6	ABM23799	Human sec	347	49	60.5	199	6	ABO38220	Human PRO
275	49	60.5	199	6	ABM29594	Human sec	348	49	60.5	199	6	ABO46130	Human PRO
276	49	60.5	199	6	ABO38525	Human sec	349	49	60.5	199	6	ABM66933	Human sec
277	49	60.5	199	6	ABO45825	Human PRO	350	49	60.5	199	6	ABD20700	Human sec
278	49	60.5	199	6	ABM20749	Human sec	351	49	60.5	199	6	ABM19834	Human sec
279	49	60.5	199	6	ADA81859	Human sec	352	49	60.5	199	6	ABO49546	Human sec
280	49	60.5	199	6	ABO16863	Human sec	353	49	60.5	199	6	ABO49851	Human sec
281	49	60.5	199	6	ABO19489	Human sec	354	49	60.5	199	6	ADA78952	Human sec
282	49	60.5	199	6	ABO22916	Human PRO	355	49	60.5	199	6	ABR88408	Human sec
283	49	60.5	199	6	ABO23221	Human PRO	356	49	60.5	199	6	ABM27154	Human sec
284	49	60.5	199	6	ABR92763	Human sec	357	49	60.5	199	6	ABM03551	Human sec
285	49	60.5	199	6	ABR81720	Human sec	358	49	60.5	199	6	ABO40050	Human sec
286	49	60.5	199	6	ABM78144	Human sec	359	49	60.5	199	7	ABO51071	Human sec
287	49	60.5	199	6	ABR89933	Human sec	360	49	60.5	199	7	ABO5527	Human sec
288	49	60.5	199	6	ABM26849	Human sec	361	49	60.5	199	7	ABO05527	Human sec
289	49	60.5	199	6	ABM13975	Human sec	362	49	60.5	199	7	ABR74831	Human sec
290	49	60.5	199	6	ABO28713	Human sec	363	49	60.5	199	7	ABR77310	Human sec
291	49	60.5	199	6	ABO30543	Human sec	364	49	60.5	199	7	ABM18067	Human sec
292	49	60.5	199	6	ABM07570	Human sec	365	49	60.5	199	7	ABR96118	Human sec
293	49	60.5	199	6	ABM04161	Human sec	366	49	60.5	199	7	ABO22051	Human sec
294	49	60.5	199	6	ABO37305	Human sec	367	49	60.5	199	7	ABO20221	Human sec
295	49	60.5	199	6	ABO41880	Human sec	368	49	60.5	199	7	ABO24524	Human sec
296	49	60.5	199	6	ABO35475	Human PRO	369	49	60.5	199	7	ABO24524	Human sec
297	49	60.5	199	6	ABM25324	Human sec	370	49	60.5	199	7	ABR86273	Human sec
298	49	60.5	199	6	ABO47716	Human sec	371	49	60.5	199	7	ABM10925	Human sec
299	49	60.5	199	6	ABO48021	Human sec	372	49	60.5	199	7	ABM76924	Human sec
300	49	60.5	199	6	ABO48631	Human sec	373	49	60.5	199	7	ABR89628	Human sec
301	49	60.5	199	6	ABO51681	Human PRO	374	49	60.5	199	7	ABM12755	Human sec
302	49	60.5	199	6	ABO51986	Human PRO	375	49	60.5	199	7	ABM06045	Human sec
303	49	60.5	199	6	ABO50766	Human sec	376	49	60.5	199	7	ABM03246	Human sec
304	49	60.5	199	6	ABR79890	Human sec	377	49	60.5	199	7	ABM19324	Human sec
305	49	60.5	199	6	ABM17152	Human sec	378	49	60.5	199	7	ABM19529	Human PRO
306	49	60.5	199	6	ABO18184	Human sec	379	49	60.5	199	7	ABO46740	Human sec
307	49	60.5	199	6	ABO21136	Human sec	380	49	60.5	199	7	ABO49241	Human sec
308	49	60.5	199	6	ABR97095	Human sec	381	49	60.5	199	7	ABR69284	Human sec
309	49	60.5	199	6	ABM12450	Human sec	382	49	60.5	199	7	ABR89323	Human sec
310	49	60.5	199	6	ABM16542	Human sec	383	49	60.5	199	7	ABR72696	Human sec
311	49	60.5	199	6	ABM24409	Human sec	384	49	60.5	199	7	ABR74526	Human sec
312	49	60.5	199	6	ABM14890	Human sec	385	49	60.5	199	7	ABO18794	Human sec
313	49	60.5	199	6	ABM04771	Human sec	386	49	60.5	199	7	ABR80500	Human sec
314	49	60.5	199	6	ABM06960	Human sec	387	49	60.5	199	7	ABM01721	Human sec
315	49	60.5	199	6	ABM09400	Human sec	388	49	60.5	199	7	ABM02331	Human sec
316	49	60.5	199	6	ABO39440	Human sec	389	49	60.5	199	7	ABR87493	Human sec

390	49	60.5	199	7	ABM13060	Abm13060 Human sec	463	49	60.5	199	8	ADJ55226	Adj55226 Human PRO
391	49	60.5	199	7	ABM30814	Abm30814 Human sec	464	49	60.5	199	8	ADJ64997	Adj64997 Human PRO
392	49	60.5	199	7	ABM24714	Abm24714 Human sec	465	49	60.5	199	8	ADM31893	Adm31893 Novel hum
393	49	60.5	199	7	ABO29628	AbO29628 Human sec	466	49	60.5	199	8	ADM36940	Adm36940 Novel hum
394	49	60.5	199	7	ABO31458	AbO31458 Human sec	467	49	60.5	199	8	ADM40745	Adm40745 Novel hum
395	49	60.5	199	7	ABM14585	Abm14585 Human sec	468	49	60.5	199	8	ADM38353	Adm38353 Novel hum
396	49	60.5	199	7	ABM10010	Abm10010 Human sec	469	49	60.5	500	4	ABB62891	Abb62891 TPO-mimet
397	49	60.5	199	7	ABO39135	AbO39135 Human sec	470	46	56.8	14	3	ABM17016	Abm17016 Drosophil
398	49	60.5	199	7	ABM34900	Abm34900 Human sec	471	46	56.8	14	5	ABB72902	Abb72902 TPO mimet
399	49	60.5	199	7	ABO51376	AbO51376 Human sec	472	46	56.8	14	8	ADJ52688	Adj52688 CHI delet
400	49	60.5	199	7	ABO04202	AbO04202 Human sec	473	46	56.8	14	8	ADJ51649	Adj51649 CHI delet
401	49	60.5	199	7	ABO10672	AbO10672 Human PRO	474	46	56.8	401	7	ABO67686	AbO67686 Kiebsiell
402	49	60.5	199	7	ABR77915	AbR77915 Human sec	475	45	55.6	14	2	AAW09484	Aaw09484 Thrombopo
403	49	60.5	199	7	ABR79125	AbR79125 Human sec	476	45	55.6	14	2	AAW36766	Aaw36766 Thrombopo
404	49	60.5	199	7	ABR24219	AbR24219 Human sec	477	45	55.6	14	2	AAW36635	Aaw36635 Thrombopo
405	49	60.5	199	7	ABR33983	AbR33983 Human sec	478	45	55.6	14	3	ABM17009	Abm17009 TPO-mimet
406	49	60.5	199	7	ABM02026	Abm02026 Human sec	479	45	55.6	14	4	AAU25854	Aau25854 Human thr
407	49	60.5	199	7	ABM78449	Abm78449 Human sec	480	45	55.6	14	5	ABB72895	Abb72895 TPO mimet
408	49	60.5	199	7	ABR90238	AbR90238 Human sec	481	45	55.6	14	7	ADJ73046	Adj73046 TPO mimet
409	49	60.5	199	7	ABM7764	Abm27764 Human sec	482	45	55.6	14	8	ADJ52681	Adj52681 CHI delet
410	49	60.5	199	7	ABM13365	Abm13365 Human sec	483	45	55.6	14	8	ADJ51642	Adj51642 CHI delet
411	49	60.5	199	7	ABO32068	AbO32068 Human sec	484	45	55.6	327	2	AAW73380	Aaw73380 CSB2 prot
412	49	60.5	199	7	ABM14280	Abm14280 Human sec	485	45	55.6	384	3	ABM37414	Abm37414 Human sec
413	49	60.5	199	7	ABM08485	Abm08485 Human sec	486	45	55.6	412	4	AAW95879	Aaw95879 Human pro
414	49	60.5	199	7	ABO40355	AbO40355 Human sec	487	45	55.6	526	2	AAW73378	Aaw73378 CSB2 prot
415	49	60.5	199	7	ABM74790	Abm74790 Human sec	488	45	55.6	526	4	AAW39051	Aaw39051 Human pol
416	49	60.5	199	7	ABM33985	Abm33985 Human sec	489	45	55.6	526	5	ABM05723	Abm05723 Human sig
417	49	60.5	199	7	ABM20444	Abm20444 Human sec	490	45	55.6	526	7	ADC37361	Adc37361 Nuclear f
418	49	60.5	199	7	ABO48936	AbO48936 Human sec	491	45	55.6	526	8	ADP55250	Adp55250 Human PRO
419	49	60.5	199	7	ABR73001	AbR73001 Human sec	492	45	55.6	558	2	AAW73379	Aaw73379 CSB2 prot
420	49	60.5	199	7	ABO15643	AbO15643 Human sec	493	44	54.3	18	7	ADN59674	Adn59674 Thrombopo
421	49	60.5	199	7	ABR85358	AbR85358 Human sec	494	44	54.3	18	7	ADN59673	Adn59673 Thrombopo
422	49	60.5	199	7	ABO15338	AbO15338 Human sec	495	44	54.3	22	7	ADN59840	Adn59840 TMP pepti
423	49	60.5	199	7	ABO17473	AbO17473 Human sec	496	44	54.3	22	7	ADN59841	Adn59841 TMP pepti
424	49	60.5	199	7	ABM17762	Abm17762 Human sec	497	44	54.3	23	7	ADN59797	Adn59797 Peptide-v
425	49	60.5	199	7	ABR85663	AbR85663 Human sec	498	44	54.3	23	7	ADN59778	Adn59778 Peptide-v
426	49	60.5	199	7	ABM77229	Abm77229 Human sec	499	44	54.3	23	7	ADN59779	Adn59779 Peptide-v
427	49	60.5	199	7	ABO28408	AbO28408 Human sec	500	44	54.3	23	7	ADN59796	Adn59796 Peptide-v
428	49	60.5	199	7	ABM33189	Abm33189 Human sec	501	44	54.3	25	7	ADN59726	Adn59726 Thrombopo
429	49	60.5	199	7	ABM30509	Abm30509 Human sec	502	44	54.3	25	7	ADN59728	Adn59728 Thrombopo
430	49	60.5	199	7	ABM21969	Abm21969 Human sec	503	44	54.3	35	7	ADN59754	Adn59754 Peptide-v
431	49	60.5	199	7	ABM16664	Abm21664 Human sec	504	44	54.3	36	7	ADN59766	Adn59766 Peptide-v
432	49	60.5	199	7	ABM15195	Abm15195 Human sec	505	44	54.3	36	7	ADN59756	Adn59756 Peptide-v
433	49	60.5	199	7	ABO41270	AbO41270 Human sec	506	44	54.3	36	7	ADN59767	Adn59767 Peptide-v
434	49	60.5	199	7	ABO37000	AbO37000 Human sec	507	44	54.3	41	7	ADN59816	Adn59816 Peptide-
435	49	60.5	199	7	ABO37610	AbO37610 Human sec	508	44	54.3	41	7	ADN59772	Adn59772 Peptide-v
436	49	60.5	199	7	ABM75400	Abm75400 Human sec	509	44	54.3	41	7	ADN59773	Adn59773 Peptide-v
437	49	60.5	199	7	ABM33680	Abm33680 Human sec	510	44	54.3	46	7	ADN59790	Adn59790 Peptide-v
438	49	60.5	199	7	ABO46435	AbO46435 Human PRO	511	44	54.3	46	7	ADN59785	Adn59785 Peptide-v
439	49	60.5	199	7	ADA83023	Ada83023 Human sec	512	44	54.3	46	7	ADN59784	Adn59784 Peptide-v
440	49	60.5	199	7	ABM32034	Abm32034 Human sec	513	44	54.3	46	7	ADN59791	Adn59791 Peptide-v
441	49	60.5	199	7	ABM31424	Abm31424 Human sec	514	44	54.3	75	7	ADN59758	Adn59758 Peptide-v
442	49	60.5	199	7	ADB86331	AdB86331 Human sec	515	44	54.3	148	8	ADO02695	Ado02695 Plant tra
443	49	60.5	199	7	ABM32339	Abm32339 Human sec	516	44	54.3	148	8	ADO02695	Ado02695 Thalecres
444	49	60.5	199	7	ABM32644	Abm32644 Human sec	517	44	54.3	148	8	ADO62315	Ado62315 Transcrip
445	49	60.5	199	7	ABM31729	Abm31729 Human sec	518	44	54.3	1039	7	ADJ70453	Adj70453 Human hea
446	49	60.5	199	7	ABM31119	Abm31119 Human sec	519	44	54.3	1098	7	ADE60625	Ade60625 Rat Prote
447	49	60.5	199	7	ADD06061	AdD06061 Human sec	520	44	54.3	1126	6	ABM07078	Abm07078 Human adi
448	49	60.5	199	7	ADG03056	AdG03056 Novel hum	521	44	54.3	1126	8	ABM0766	Abm0766 Tumour-as
449	49	60.5	199	7	ADG01763	AdG01763 Novel hum	522	44	54.3	1126	8	ABM0767	Abm0767 Tumour-as
450	49	60.5	199	7	ADF95938	AdF95938 Novel hum	523	44	54.3	1132	7	ADE60627	Ade60627 Human PRO
451	49	60.5	199	7	ADG12753	AdG12753 Novel hum	524	44	54.3	1132	7	ADE60623	Ade60623 Human PRO
452	49	60.5	199	7	ADH09413	AdH09413 Human PRO	525	44	54.3	1132	8	ABM0765	Abm0765 Tumour-as
453	49	60.5	199	7	ADL33192	AdL33192 Novel hum	526	43	53.1	18	7	ADN59678	Adn59678 Thrombopo
454	49	60.5	199	7	ADM30728	Adm30728 Novel hum	527	43	53.1	18	7	ADN59677	Adn59677 Thrombopo
455	49	60.5	199	8	ADE74725	Ade74725 Human sec	528	43	53.1	25	7	ADN59734	Adn59734 Thrombopo
456	49	60.5	199	8	ADE75337	Ade75337 Human sec	529	43	53.1	25	7	ADN59736	Adn59736 Thrombopo
457	49	60.5	199	8	ADP96550	AdP96550 Novel hum	530	43	53.1	62	3	ABA42889	AbA42889 Human ORF
458	49	60.5	199	8	ADG04821	AdG04821 Novel hum	531	42	51.9	14	4	AAW98500	Aaw98500 Human pep
459	49	60.5	199	8	ADG00981	AdG00981 Novel hum	532	42	51.9	29	2	AAO28894	Aay028894 Fragment
460	49	60.5	199	8	ADG83237	AdG83237 Human PRO	533	42	51.9	29	7	ADA07724	Ada07724 Human sec
461	49	60.5	199	8	ADH26518	AdH26518 Novel hum	534	42	51.9	29	8	ADN41333	Adn41333 Novel hum
462	49	60.5	199	8	ADH33487	Adh33487 Human PRO	535	42	51.9	43	3	AAG57089	Aag57089 Arabidops

536	42	51.9	93	4	AAU50884	Propionib	Aau50884	609	41	50.6	126	8	ADO10073	Novel hum
537	42	51.9	93	6	ABM47403	Propionib	Abm47403	610	41	50.6	159	8	ADU02995	Novel hum
538	42	51.9	100	2	AAU59831	Human nor	Aay59831	611	41	50.6	173	8	ADU02996	Novel hum
539	42	51.9	143	8	ADL04815	M. catarr	Adl04815	612	41	50.6	203	6	ABP99502	Human sec
540	42	51.9	171	8	ABO58540	Human gen	Abo58540	613	41	50.6	203	6	ABR00975	Human gen
541	42	51.9	171	8	ABO58663	Human gen	Abo58663	614	41	50.6	203	6	ABR00975	Human gen
542	42	51.9	189	5	AAO17523	Human M32	Aao17523	615	41	50.6	203	6	ADA44031	Human sec
543	42	51.9	196	7	ADC32882	Human nov	Adc32882	616	41	50.6	203	7	ADC20347	Human sec
544	42	51.9	216	3	ABAB12489	Human PSE	Abab12489	617	41	50.6	203	7	ADF10687	Human sec
545	42	51.9	225	4	AAU25636	G protein	Aau25636	618	41	50.6	204	3	AAB51840	Human sec
546	42	51.9	300	7	ADC86753	Human GPC	Adc86753	619	41	50.6	205	3	ABAB32114	Human sec
547	42	51.9	320	4	ABO07921	Murine pe	Abp07921	620	41	50.6	320	4	ABAB4502	Human pro
548	42	51.9	419	5	ABP07921	Murine pe	Abp07921	621	41	50.6	320	4	ABP07921	Murine pe
549	42	51.9	419	5	ABP07921	Murine pe	Abp07921	622	41	50.6	320	5	AAO17516	Human M30
550	42	51.9	420	5	ABP07921	Murine pe	Abp07921	623	41	50.6	320	5	AAO17516	Human M30
551	42	51.9	420	5	ABP07921	Murine pe	Abp07921	624	41	50.6	320	5	AAO17516	Human M30
552	42	51.9	420	6	ABR41083	Human nov	AbR41083	625	41	50.6	320	5	AAO17516	Human M30
553	42	51.9	420	7	ADC31180	Human nov	Adc31180	626	41	50.6	320	5	AAO17516	Human M30
554	42	51.9	420	8	ADP48671	Human pel	Adp48671	627	41	50.6	320	5	AAO17516	Human M30
555	42	51.9	420	8	ABAB95188	Human pro	Abab95188	628	41	50.6	320	5	AAO17516	Human M30
556	42	51.9	420	8	ABAB8569	Human pro	Abab8569	629	41	50.6	320	5	AAO17516	Human M30
557	42	51.9	487	4	ABAB8569	Human pro	Abab8569	630	41	50.6	320	5	AAO17516	Human M30
558	42	51.9	487	6	ABJ20215	Human SLC	Abj20215	631	41	50.6	320	5	AAO17516	Human M30
559	42	51.9	487	6	ABU04694	Human exp	Abu04694	632	41	50.6	320	5	AAO17516	Human M30
560	42	51.9	487	6	ABU04694	Human exp	Abu04694	633	41	50.6	320	5	AAO17516	Human M30
561	42	51.9	487	6	ABU04694	Human exp	Abu04694	634	41	50.6	320	5	AAO17516	Human M30
562	42	51.9	487	6	ABU04694	Human exp	Abu04694	635	41	50.6	320	5	AAO17516	Human M30
563	42	51.9	487	6	ABU04694	Human exp	Abu04694	636	41	50.6	320	5	AAO17516	Human M30
564	42	51.9	487	6	ABU04694	Human exp	Abu04694	637	41	50.6	320	5	AAO17516	Human M30
565	42	51.9	487	6	ABU04694	Human exp	Abu04694	638	41	50.6	320	5	AAO17516	Human M30
566	42	51.9	487	6	ABU04694	Human exp	Abu04694	639	41	50.6	320	5	AAO17516	Human M30
567	42	51.9	487	6	ABU04694	Human exp	Abu04694	640	41	50.6	320	5	AAO17516	Human M30
568	42	51.9	487	6	ABU04694	Human exp	Abu04694	641	41	50.6	320	5	AAO17516	Human M30
569	42	51.9	487	6	ABU04694	Human exp	Abu04694	642	41	50.6	320	5	AAO17516	Human M30
570	42	51.9	487	6	ABU04694	Human exp	Abu04694	643	41	50.6	320	5	AAO17516	Human M30
571	42	51.9	487	6	ABU04694	Human exp	Abu04694	644	41	50.6	320	5	AAO17516	Human M30
572	42	51.9	487	6	ABU04694	Human exp	Abu04694	645	41	50.6	320	5	AAO17516	Human M30
573	42	51.9	487	6	ABU04694	Human exp	Abu04694	646	41	50.6	320	5	AAO17516	Human M30
574	42	51.9	487	6	ABU04694	Human exp	Abu04694	647	41	50.6	320	5	AAO17516	Human M30
575	42	51.9	487	6	ABU04694	Human exp	Abu04694	648	41	50.6	320	5	AAO17516	Human M30
576	42	51.9	487	6	ABU04694	Human exp	Abu04694	649	41	50.6	320	5	AAO17516	Human M30
577	42	51.9	487	6	ABU04694	Human exp	Abu04694	650	41	50.6	320	5	AAO17516	Human M30
578	42	51.9	487	6	ABU04694	Human exp	Abu04694	651	41	50.6	320	5	AAO17516	Human M30
579	42	51.9	487	6	ABU04694	Human exp	Abu04694	652	41	50.6	320	5	AAO17516	Human M30
580	41.5	51.2	18	7	ADN59681	Thrombopo	Adn59681	653	40	49.4	14	2	AAW09485	Thrombopo
581	41.5	51.2	25	7	ADN59742	Thrombopo	Adn59742	654	40	49.4	14	2	AAW09485	Thrombopo
582	41.5	51.2	43	7	ADN59752	Peptide-v	Adn59752	655	40	49.4	14	2	AAW09485	Thrombopo
583	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	656	40	49.4	14	2	AAW09485	Thrombopo
584	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	657	40	49.4	14	2	AAW09485	Thrombopo
585	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	658	40	49.4	14	2	AAW09485	Thrombopo
586	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	659	40	49.4	14	2	AAW09485	Thrombopo
587	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	660	40	49.4	14	2	AAW09485	Thrombopo
588	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	661	40	49.4	14	2	AAW09485	Thrombopo
589	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	662	40	49.4	14	2	AAW09485	Thrombopo
590	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	663	40	49.4	14	2	AAW09485	Thrombopo
591	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	664	40	49.4	14	2	AAW09485	Thrombopo
592	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	665	40	49.4	14	2	AAW09485	Thrombopo
593	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	666	40	49.4	14	2	AAW09485	Thrombopo
594	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	667	40	49.4	14	2	AAW09485	Thrombopo
595	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	668	40	49.4	14	2	AAW09485	Thrombopo
596	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	669	40	49.4	14	2	AAW09485	Thrombopo
597	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	670	40	49.4	14	2	AAW09485	Thrombopo
598	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	671	40	49.4	14	2	AAW09485	Thrombopo
599	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	672	40	49.4	14	2	AAW09485	Thrombopo
600	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	673	40	49.4	14	2	AAW09485	Thrombopo
601	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	674	40	49.4	14	2	AAW09485	Thrombopo
602	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	675	40	49.4	14	2	AAW09485	Thrombopo
603	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	676	40	49.4	14	2	AAW09485	Thrombopo
604	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	677	40	49.4	14	2	AAW09485	Thrombopo
605	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	678	40	49.4	14	2	AAW09485	Thrombopo
606	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	679	40	49.4	14	2	AAW09485	Thrombopo
607	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	680	40	49.4	14	2	AAW09485	Thrombopo
608	41	50.6	13	3	ABAB17015	TPO-mimet	Abab17015	681	40	49.4	14	2	AAW09485	Thrombopo

682	40	49.4	406	7	ADB65774	Adb65774 Human pro	755	38	46.9	105	6	ABU55355	Abu55355 Human nov
683	40	49.4	409	2	AAW26526	Aaw26526 Branched	756	38	46.9	141	6	ABM72249	Abm72249 Staphyloc
684	40	49.4	409	4	ABAB61179	Abab61179 Pseudomon	757	38	46.9	151	4	AAU44520	Aau44520 Propionib
685	40	49.4	410	6	ABU40351	Abu40351 Protein e	758	38	46.9	151	6	ABM41039	Abm41039 Propionib
686	40	49.4	415	8	ADO44009	Ado44009 Amino aci	759	38	46.9	154	8	ADQ66223	Adq66223 Novel hum
687	40	49.4	419	6	ABR41082	Abm41082 Mouse MAP	760	38	46.9	158	4	AAW39969	Aaw39969 Human pol
688	40	49.4	419	8	ADQ67161	Adq67161 Novel hum	761	38	46.9	166	2	AAW94848	Aaw94848 Human cer
689	40	49.4	445	5	ABB07923	Abb07923 Human pel	762	38	46.9	166	8	ADR08617	Adr08617 Human pro
690	40	49.4	445	5	AAO17520	Aao17520 Murine M3	763	38	46.9	176	8	ADX93119	Adx93119 Plant ful
691	40	49.4	445	5	AAO17521	Aao17521 Human M31	764	38	46.9	177	4	AAE03109	Aae03109 Human gen
692	40	49.4	445	7	ADC31807	Adc31807 Human nov	765	38	46.9	177	5	ABG63620	Abg63620 Human alb
693	40	49.4	445	7	ADD71147	Add71147 Human int	766	38	46.9	177	8	ADL76885	Adl76885 Alhuman f
694	40	49.4	445	8	ADP48675	Adp48675 Human pel	767	38	46.9	180	6	ABP96348	Abp96348 Human AGE
695	40	49.4	464	7	ABO67914	Abob67914 Pseudomon	768	38	46.9	181	9	ABM95454	Abm95454 M. xanthu
696	40	49.4	476	8	ADS42549	Ads42549 Bacterial	769	38	46.9	186	8	ADX97029	Adx97029 Plant ful
697	40	49.4	506	6	ADA55185	Ada55185 Human pro	770	38	46.9	206	4	AAW41755	Aaw41755 Human pol
698	40	49.4	554	7	ADM04654	Adm04654 Human pro	771	38	46.9	206	9	ADX08152	Adx08152 Cyclin-de
699	40	49.4	561	8	ADN27155	Adn27155 Bacterial	772	38	46.9	219	2	AAW24021	Aaw24021 Human pho
700	40	49.4	715	9	ABM93740	Abm93740 M. xanthu	773	38	46.9	219	9	ADX08150	Adx08150 Cyclin-de
701	40	49.4	881	5	ADH48730	Adh48730 NOV6 prot	774	38	46.9	235	8	ADX75054	Adx75054 Plant ful
702	40	49.4	881	8	ADU02359	Adu02359 Novel hum	775	38	46.9	264	4	AAW71897	Aaw71897 Human olf
703	40	49.4	1068	7	ADC39124	Adc39124 Novel hum	776	38	46.9	267	2	AAW91362	Aaw91362 Proenkeph
704	40	49.4	1077	4	AAU01924	Aau01924 Human ade	777	38	46.9	267	4	AAW89159	Aaw89159 Human sec
705	40	49.4	1077	8	ADL14164	Adl14164 Novel hum	778	38	46.9	267	8	ABG75474	Abg75474 Human bet
706	40	49.4	1086	5	AAE17129	Aae17129 Human ade	779	38	46.9	267	8	ABG75473	Abg75473 Human pro
707	40	49.4	1103	5	ABB81465	Abb81465 Human ade	780	38	46.9	289	5	AAU95669	Aau95669 Human olf
708	39	48.1	58	8	ABO56054	Abob56054 Human gen	781	38	46.9	289	8	ADG83416	Adg83416 Human olf
709	39	48.1	76	4	AAU61033	Aau61033 Propionib	782	38	46.9	297	3	AAW2752	Aaw2752 Arabidops
710	39	48.1	76	6	ABM57552	Abm57552 Propionib	783	38	46.9	297	3	ABG16893	Abg16893 Arabidops
711	39	48.1	77	4	AAU53862	Aau53862 Propionib	784	38	46.9	297	4	ABM11849	Abm11849 Human pre
712	39	48.1	77	6	ABM50381	Abm50381 Propionib	785	38	46.9	313	6	ABR01653	Abro1653 Human G p
713	39	48.1	82	6	ABP58654	Abp58654 Human mil	786	38	46.9	319	6	ABR01638	Abro1638 Human G p
714	39	48.1	260	6	ABP58622	Abp58622 Human pro	787	38	46.9	320	5	ABG76863	Abg76863 Human G-p
715	39	48.1	260	6	ABO00904	Abob00904 Polypepti	788	38	46.9	320	5	ABP95882	Abp95882 Human GPC
716	39	48.1	308	5	AAU95592	Aau95592 Human olf	789	38	46.9	320	5	AAU95784	Aau95784 Human olf
717	39	48.1	308	6	ABO00671	Abob00671 Novel hum	790	38	46.9	320	5	AAU95782	Aau95782 Human G p
718	39	48.1	324	4	AAW71587	Aaw71587 Human olf	791	38	46.9	320	6	ABP97086	Abp97086 Human G p
719	39	48.1	324	5	ABP51609	Abp51609 Human G p	792	38	46.9	320	6	ABP97085	Abp97085 Human G p
720	39	48.1	324	7	AD104343	Adi04343 Human G-p	793	38	46.9	320	7	ADC86249	Adc86249 Human GPC
721	39	48.1	335	4	AAU24741	Aau24741 Human olf	794	38	46.9	320	7	ADC86247	Adc86247 Human GPC
722	39	48.1	335	5	ABB06604	Abb06604 G protein	795	38	46.9	320	8	ADG83532	Adg83532 Human Olf
723	39	48.1	335	5	ABB06605	Abb06605 G protein	796	38	46.9	320	8	ADG83536	Adg83536 Human Olf
724	39	48.1	335	5	ABB06606	Abb06606 G protein	797	38	46.9	333	3	AAW16892	Aaw16892 Arabidops
725	39	48.1	335	5	ABP95761	Abp95761 Human GPC	798	38	46.9	333	3	AAW52751	Aaw52751 Arabidops
726	39	48.1	335	5	AAU85361	Aau85361 G-coupled	799	38	46.9	337	3	AAW52750	Aaw52750 Arabidops
727	39	48.1	335	5	ADC85603	Adc85603 Human GPC	800	38	46.9	337	3	AAW16891	Aaw16891 Arabidops
728	39	48.1	341	3	AAW69880	Aaw69880 B. lactof	801	38	46.9	368	4	ABG03113	Abg03113 Novel hum
729	39	48.1	379	8	ADP74528	Adp74528 Plant ful	802	38	46.9	426	7	ABO62034	Abob62034 Klebsiell
730	39	48.1	460	8	ADP79710	Adp79710 Cycloclas	803	38	46.9	433	4	ABG04452	Abg04452 Novel hum
731	39	48.1	461	9	ABM94829	Abm94829 M. xanthu	804	38	46.9	433	7	ADE09953	Ade09953 Novel pro
732	39	48.1	491	7	ADG84110	Adg84110 Human TMD	805	38	46.9	433	9	AEA21042	Aea21042 Novel hum
733	39	48.1	517	4	ABB62049	Abb62049 Drosophil	806	38	46.9	435	8	ADY11835	Ady11835 Plant ful
734	39	48.1	716	6	ABU33863	Abu33863 Protein e	807	38	46.9	447	7	ADC31461	Adc31461 Human nov
735	39	48.1	966	6	ABU56300	Abu56300 Caspase r	808	38	46.9	450	4	AAU56552	Aau56552 Propionib
736	39	48.1	1086	6	ABU46534	Abu46534 Protein e	809	38	46.9	450	6	ABM53071	Abm53071 Propionib
737	39	48.1	1086	6	ABE91670	Abep91670 Microbial	810	38	46.9	454	3	AAW53151	Aaw53151 Macaca mu
738	39	48.1	1525	7	ABM89427	Abm89427 Rice abio	811	38	46.9	476	8	ADT58335	Adt58335 Plant pol
739	38	47.5	265	7	ABO63876	Abob63876 Klebsiell	812	38	46.9	493	5	ABW93470	Abw93470 Herbicida
740	38	46.9	22	8	ADG20872	Adg20872 TMII-VII	813	38	46.9	524	5	ABW65112	Abw65112 Human NOV
741	38	46.9	36	4	ABB39297	Abb39297 Peptide #	814	38	46.9	524	8	ADN61875	Adn61875 Human nov
742	38	46.9	36	4	ABM32805	Abm32805 Peptide #	815	38	46.9	526	8	ADY07792	Ady07792 Plant ful
743	38	46.9	36	4	AAW72565	Aaw72565 Human bon	816	38	46.9	530	5	ABW65113	Abw65113 Human NOV
744	38	46.9	36	4	AAW59965	Aaw59965 Human bra	817	38	46.9	530	8	ADN61877	Adn61877 Human nov
745	38	46.9	36	4	ABG54257	Abg54257 Human liv	818	38	46.9	545	7	ABO80203	Abob80203 Pseudomon
746	38	46.9	36	5	ABG42386	Abg42386 Human pep	819	38	46.9	549	3	AAW73346	Aaw73346 HTSM clon
747	38	46.9	37	3	AAW34113	Aaw34113 Human sec	820	38	46.9	577	4	ABG13760	Abg13760 Novel hum
748	38	46.9	84	5	AAE20819	Aae20819 Human gen	821	38	46.9	580	3	AAW43912	Aaw43912 Human can
749	38	46.9	84	5	ABG64683	Abg64683 Human alb	822	38	46.9	580	7	ADP55450	Adp55450 Human nov
750	38	46.9	84	8	ADL77950	Adl77950 Alhuman f	823	38	46.9	603	4	ABG00654	Abg00654 Novel hum
751	38	46.9	105	8	ADX92954	Adx92954 Plant ful	824	38	46.9	603	4	ABG08209	Abg08209 Novel hum
752	38	46.9	105	4	AAU16286	Aau16286 Human nov	825	38	46.9	739	4	ABG03917	Abg03917 Novel hum
753	38	46.9	105	4	AAU29502	Aau29502 Human G p	826	38	46.9	781	4	ABG24696	Abg24696 Novel hum
754	38	46.9	105	5	ABG60790	Abg60790 Novel G p	827	38	46.9	790	6	ADA55066	Ada55066 Human pro

828	38	46.9	790	7	ADC31848	Adc31848 Human nov	901	37	45.7	61	6	ABP76329	Abp76329 Human GEN
829	38	46.9	790	7	ADJ71094	Adj71094 Human hea	902	37	45.7	65	4	AAU63303	Aau63303 Propionib
830	38	46.9	855	8	ADN74169	Adn74169 Thale cre	903	37	45.7	65	6	ABM59822	Abm59822 Propionib
831	38	46.9	994	8	ADQ33393	Adq33393 Human myo	904	37	45.7	72	4	AAU49140	Aau49140 Propionib
832	38	46.9	994	8	ADQ33395	Adq33395 Human myo	905	37	45.7	72	6	ABM45659	Abm45659 Propionib
833	38	46.9	994	8	ADQ33396	Adq33396 Human myo	906	37	45.7	91	4	ABBI5848	Abbi5848 Human ner
834	38	46.9	1027	8	ADQ33394	Adq33394 Human myo	907	37	45.7	103	6	ADA55455	Ada55455 Human pro
835	38	46.9	1081	7	ADE40146	Ade40146 Human NOV	908	37	45.7	107	3	AAB41227	Aab41227 Human ORF
836	38	46.9	1116	5	AAU97582	Aau97582 Human CD1	909	37	45.7	107	5	ADL92620	Adl92620 IMABIS003
837	38	46.9	1116	5	ADQ17792	Adq17792 Human sof	910	37	45.7	110	8	ADR41966	Adr41966 VAP relat
838	38	46.9	1116	9	ADV70266	Adv70266 Tumorr-ass	911	37	45.7	110	9	ADX70284	Adx70284 Versatile
839	38	46.9	1121	4	AAU39493	Aau39493 Human pol	912	37	45.7	111	4	ABG02805	Abg02805 Novel hum
840	38	46.9	1124	4	AAU41280	Aau41280 Human pol	913	37	45.7	123	2	AAU76599	Aau76599 Human ova
841	38	46.9	1124	4	AAU41279	Aau41279 Human pol	914	37	45.7	123	2	ABO76599	Abu76599 Pseudomon
842	38	46.9	1149	5	AAU97583	Aau97583 Human CD1	915	37	45.7	140	7	ABO80854	Abu80854 Pseudomon
843	38	46.9	1151	5	AAU97585	Aau97585 Human CD1	916	37	45.7	143	4	AAU14876	Aau14876 Peptide #
844	38	46.9	1154	4	AAU39494	Aau39494 Human pol	917	37	45.7	143	4	AAU14879	Aau14879 Peptide #
845	38	46.9	1156	5	AAU97584	Aau97584 Human CD1	918	37	45.7	143	4	ABB33848	Abb33848 Peptide #
846	38	46.9	1156	8	ADQ33397	Adq33397 Human myo	919	37	45.7	143	4	ABB33845	Abb33845 Peptide #
847	38	46.9	1232	8	ADE45189	Ade45189 Human CD1	920	37	45.7	143	4	ABM27305	Abm27305 Peptide #
848	38	46.9	1492	6	AAU31008	Aau31008 Human tra	921	37	45.7	143	4	AAU27308	Aau27308 Peptide #
849	38	46.9	1492	7	ADM29378	Adm29378 Human nov	922	37	45.7	143	4	ABB28663	Abb28663 Peptide #
850	38	46.9	1492	7	ADM29378	Adm29378 Human nov	923	37	45.7	143	4	ABB28661	Abb28661 Peptide #
851	38	46.9	1492	7	ADM29378	Adm29378 Human nov	924	37	45.7	143	4	ABB19289	Abb19289 Protein #
852	38	46.9	1545	2	AAU33362	Aau33362 Human can	925	37	45.7	143	4	ABB19287	Abb19287 Protein #
853	38	46.9	1545	2	AAU33362	Aau33362 Human can	926	37	45.7	143	4	ABB19287	Abb19287 Protein #
854	38	46.9	1545	5	AAU29040	Aau29040 Human cMO	927	37	45.7	143	4	AAU67018	Aau67018 Human bon
855	38	46.9	1545	5	AAU29040	Aau29040 Human cMO	928	37	45.7	143	4	AAU67018	Aau67018 Human bon
856	38	46.9	1545	8	ADL60256	Adl60256 Human org	929	37	45.7	143	4	AAU54610	Aau54610 Human bra
857	38	46.9	1545	8	ADL60270	Adl60270 Human org	930	37	45.7	143	4	AAU54612	Aau54612 Human bra
858	38	46.9	1545	8	ADL60272	Adl60272 Human org	931	37	45.7	143	4	ABG48681	Abg48681 Human liv
859	38	46.9	1545	8	ADL60268	Adl60268 Human org	932	37	45.7	143	4	ABG48683	Abg48683 Human liv
860	38	46.9	1545	8	ADL60250	Adl60250 Human org	933	37	45.7	143	4	AAU02601	Aau02601 Peptide #
861	38	46.9	1545	8	ADL60258	Adl60258 Human org	934	37	45.7	143	5	ABG36675	Abg36675 Human pep
862	38	46.9	1545	8	ADL60248	Adl60248 Human org	935	37	45.7	143	5	ABG36673	Abg36673 Human pep
863	38	46.9	1545	8	ADL60248	Adl60248 Human org	936	37	45.7	143	8	ABO59053	Abu59053 Human gen
864	38	46.9	1545	8	ADL60228	Adl60228 Human org	937	37	45.7	147	8	ADR09851	Adr09851 Human pro
865	38	46.9	1545	8	ADL60252	Adl60252 Human org	938	37	45.7	152	3	AAU59906	Aau59906 Arabidops
866	38	46.9	1545	8	ADL60254	Adl60254 Human org	939	37	45.7	177	4	ABG28550	Abg28550 Novel hum
867	38	46.9	1545	8	ADL60260	Adl60260 Human org	940	37	45.7	205	4	ABG28550	Abg28550 Novel hum
868	38	46.9	1545	8	ADL60262	Adl60262 Human org	941	37	45.7	222	4	ABG10784	Abg10784 Novel hum
869	38	46.9	1545	8	ADL60266	Adl60266 Human org	942	37	45.7	248	5	AAU91164	Aau91164 Pantothan
870	38	46.9	1545	8	ADP20017	Adp20017 Human can	943	37	45.7	249	8	ADS21323	Ads21323 Bacterial
871	37.5	46.3	198	9	ADP20017	Adp20017 Human can	944	37	45.7	253	7	ABO72591	Abu72591 Pseudomon
872	37.5	46.3	509	3	ABU13371	Abu13371 Human can	945	37	45.7	266	7	ADC85016	Adc85016 HLA DRB1
873	37	45.7	15	5	ABU13358	Abu13358 Streptomy	946	37	45.7	266	7	ADC85013	Adc85013 HLA DRB1
874	37	45.7	15	5	ABU13358	Abu13358 Streptomy	947	37	45.7	266	7	ADC85019	Adc85019 HLA DRB1
875	37	45.7	15	5	ABU13358	Abu13358 Streptomy	948	37	45.7	267	8	ABO60229	Abu60229 Human gen
876	37	45.7	15	8	ADJ52691	Adj52691 TPO-mimet	949	37	45.7	272	3	AAU06655	Aau06655 Arabidops
877	37	45.7	15	8	ADJ52690	Adj52690 CH1 delet	950	37	45.7	281	3	AAU06654	Aau06654 Arabidops
878	37	45.7	15	8	ADJ52690	Adj52690 CH1 delet	951	37	45.7	281	3	AAU06654	Aau06654 Arabidops
879	37	45.7	15	8	ADJ51652	Adj51652 CH1 delet	952	37	45.7	282	3	AAU060574	Aau060574 Arabidops
880	37	45.7	15	8	ADJ51651	Adj51651 CH1 delet	953	37	45.7	282	3	AAU060574	Aau060574 Arabidops
881	37	45.7	18	2	AAU09460	Aau09460 Thrombopo	954	37	45.7	282	3	AAU060574	Aau060574 Arabidops
882	37	45.7	18	2	AAU09460	Aau09460 Thrombopo	955	37	45.7	284	5	ABU93853	Abu93853 Arabidops
883	37	45.7	18	2	AAU09460	Aau09460 Thrombopo	956	37	45.7	289	4	ABU93853	Abu93853 Arabidops
884	37	45.7	18	2	AAU09460	Aau09460 Thrombopo	957	37	45.7	289	4	ABU93853	Abu93853 Arabidops
885	37	45.7	18	4	AAU25824	Aau25824 Human thr	958	37	45.7	310	4	AAU24739	Aau24739 Human olf
886	37	45.7	18	4	AAU25824	Aau25824 Human thr	959	37	45.7	310	5	ABU76839	Abu76839 Human G-p
887	37	45.7	18	4	AAU25824	Aau25824 Human thr	960	37	45.7	310	5	ABU76839	Abu76839 Human G-p
888	37	45.7	18	5	ABU72912	Abu72912 TPO mimet	961	37	45.7	310	5	ABU76839	Abu76839 Human G-p
889	37	45.7	18	7	ADN73064	Adn73064 TPO mimet	962	37	45.7	310	5	AAU95691	Aau95691 Human olf
890	37	45.7	18	7	ADN73064	Adn73064 TPO mimet	963	37	45.7	310	5	AAU95691	Aau95691 Human olf
891	37	45.7	18	8	ADJ52699	Adj52699 CH1 delet	964	37	45.7	310	6	ABU11155	Abu11155 Human G-p
892	37	45.7	18	8	ADJ52699	Adj52699 CH1 delet	965	37	45.7	310	7	ADU04319	Adu04319 Human G-p
893	37	45.7	22	7	ADN59775	Adn59775 Peptide-v	966	37	45.7	319	3	AAU90513	Aau90513 Rat cereb
894	37	45.7	23	7	ADN59775	Adn59775 Peptide-v	967	37	45.7	319	3	AAU90513	Aau90513 Rat cereb
895	37	45.7	23	7	ADN59775	Adn59775 Peptide-v	968	37	45.7	319	8	ABO58539	Abu58539 Human gen
896	37	45.7	25	7	ADN59696	Adn59696 Thrombopo	969	37	45.7	344	2	AAU30208	Aau30208 C. albica
897	37	45.7	36	7	ADN59763	Adn59763 Peptide-v	970	37	45.7	344	2	AAU30208	Aau30208 C. albica
898	37	45.7	36	7	ADN59763	Adn59763 Peptide-v	971	37	45.7	344	2	AAU30208	Aau30208 C. albica
899	37	45.7	46	7	ADN59781	Adn59781 Peptide-v	972	37	45.7	345	5	ABP73737	Abp73737 Candida a
900	37	45.7	46	7	ADN59787	Adn59787 Peptide-v	973	37	45.7	375	7	ADC94127	Adc94127 E. faeciu

974 37 45.7 410 6 ABU22259 Protein e
 975 37 45.7 410 6 ABU20206 Protein e
 976 37 45.7 410 9 AEB38098 L. pneumo
 977 37 45.7 410 9 AEB41389 L. pneumo
 978 37 45.7 416 4 AAU04813 Micromono
 979 37 45.7 416 6 ABP99322 Orthosomy
 980 37 45.7 422 7 ABO60778 Klebsiell
 981 37 45.7 427 9 AEB42134 L. pneumo
 982 37 45.7 431 9 AEB38990 L. pneumo
 983 37 45.7 435 4 AAM39070 Human pol
 984 37 45.7 445 3 AAG60569 Arabidops
 985 37 45.7 449 7 ABO68161 Pseudomon
 986 37 45.7 451 3 AAG60573 Arabidops
 987 37 45.7 451 3 AAG20634 Arabidops
 988 37 45.7 452 3 AAB23474 Burkholde
 989 37 45.7 454 2 AAR97842 Kaposi's
 990 37 45.7 454 2 AAR93616 Kaposi's
 991 37 45.7 457 7 ADM04739 Human pro
 992 37 45.7 458 9 AEB28955 Anopheles
 993 37 45.7 483 4 AAM40856 Human pol
 994 37 45.7 504 5 ABB89686 Human pol
 995 37 45.7 513 4 ABB65614 Drosophil
 996 37 45.7 538 4 ABG21068 Novel hum
 997 37 45.7 566 4 ABB61743 Drosophil
 998 37 45.7 575 5 ABP27731 Streptoco
 999 37 45.7 625 5 ABB81615 Human HDA
 1000 37 45.7 636 4 AAE05105 Drosophil

ALIGNMENTS

RESULT 1
 AAU09483
 ID AAU09483 standard; protein; 14 AA.

AC AAU09483;
 DT 10-SEP-1997 (first entry)
 XX

Thrombopoietin receptor binding peptide.

Heematology; thrombocytopenia; TPO; TR; proliferation;
 bone marrow transfusion; chemotherapy; radiation therapy.

Synthetic.

WO9640189-A1.

19-DEC-1996.

05-JUN-1996; 96WO-US008998.

07-JUN-1995; 95US-00472371.

07-JUN-1995; 95US-00473604.

07-JUN-1995; 95US-00476168.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00484090.

07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-051883/05.

Thrombopoietin receptor-binding/activating peptide(s) and peptide

mimetic(s) - useful in treatment of haematological disorders, esp.

thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 26; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells
 XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CELVGPSTLMSWLTC 14

Db 1 CELVGPSTLMSWLTC 14

RESULT 2

AAW36634

ID AAW36634 standard; peptide; 14 AA.

XX

AAW36634;

DT 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;

KW haematological disorder; thrombocytopenia; chemotherapy;

KW radiation therapy; bone marrow transfusion; diagnosis;

KW signal transduction; receptor activation; cell culture.

OS Synthetic.

WO9640750-A1.

19-DEC-1996.

07-JUN-1996; 96WO-US009623.

07-JUN-1995; 95US-00478128.

07-JUN-1995; 95US-00485301.

(GLAX) GLAXO GROUP LTD.

Dower WJ, Barrett RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;

Matheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

WPI; 1997-052226/05.

Peptides and peptide mimetics which bind to and activate the

thrombopoietin receptor - useful in treatment of haematological

disorders, esp. thrombocytopenia resulting from chemotherapy, etc.

Disclosure; Page 26; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TR), can be

used to treat disorders which are susceptible to treatment with a

thrombopoietin agonist, preferably haematological disorders and

thrombocytopenia resulting from chemotherapy, radiation therapy or bone

marrow transfusions. It can also be used diagnostically, e.g. to

investigate the mechanism of thrombopoietin signal transduction and

receptor activation, or to maintain the proliferation and growth of

thrombopoietin dependent cell lines

XX Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;

Best Local Similarity 100.0%; Pred. NO. 1.5e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPMSLWLTTC 14
 DB 1 CELVGPMSLWLTTC 14

RESULT 3

AAW36767
 ID AAW36767 standard; peptide; 14 AA.

XX AAW36767;

DT 11-MAR-1998 (first entry)

XX Thrombopoietin receptor binding peptide.

XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.

XX Synthetic.

XX Key Location/Qualifiers

XX Disulfide-bond 1. 14

XX Modified-site 14

XX FT /note= "NH2-Cys"

XX W09640750-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US009623.

XX 07-JUN-1995; 95US-00478128.

XX 07-JUN-1995; 95US-00485301.

XX (GLAXO) GLAXO GROUP LTD.

XX Dower WJ, Barret RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;

PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-052226/05.

XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.

XX Example 9; Page 75; 106pp; English.

XX The present peptide, which binds the thrombopoietin receptor (TP), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines

SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPMSLWLTTC 14

DB 1 CELVGPMSLWLTTC 14

RESULT 4

AAAB17011
 ID AAB17011 standard; peptide; 14 AA.

XX

AC AAB17011;

XX 31-OCT-2000 (first entry)

DE TPO-mimetic peptide sequence SEQ ID NO:67.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.

XX Synthetic.

XX W0200024782-A2.

XX 04-MAY-2000.

XX 25-OCT-1999; 99WO-US025044.

XX 23-OCT-1998; 98US-0105371P.

XX 22-OCT-1999; 99US-00428082.

XX (AMGE-) AMGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and pharmacologically
 active peptides, useful for treating cancer and autoimmune diseases.

XX Claim 19; Page 217; 608pp; English.

XX The present invention describes composition of matter (1) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (1) is:
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently 0 or 1, provided that at least 1
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AA69443 to AA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention

SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPMSLWLTTC 14

DB 1 CELVGPMSLWLTTC 14

RESULT 5

AAU25853
 ID AAU25853 standard; peptide; 14 AA.

XX AAU25853;

XX 17-DEC-2001 (first entry)

XX DE Human thrombopoietin receptor (TPO-R) activator peptide #39.

XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.

XX OS Homo sapiens.

XX KW US6251864-B1.

XX PD 26-JUN-2001.

XX PF 01-MAR-2000; 2000US-00516704.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00485301.

XX PR 07-JUN-1996; 96WO-US0009623.

XX PR 15-AUG-1996; 96US-00699027.

XX PA (GLAXO) GLAXO GROUP LTD.

XX PI Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
 PI Yin Q;

XX DR MPI; 2001-564142/63.

XX PT Activating thrombopoietin receptors in cells, used to treat
 PT thrombocytopenia and hematological disorders, comprises contacting cells
 PT with peptides and peptide mimetics attached to hydrophilic polymers.

XX PS Disclosure; Col 20; 128pp; English.

XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent hematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWLTTC 14
 |||||
 Db 1 CELVGPMSLMSWLTTC 14

RESULT 6
 ABB72897
 ID ABB72897 standard; peptide; 14 AA.
 XX
 AC ABB72897;
 XX

DT XX TPO mimetic peptide SEQ ID NO:67.

DE XX

XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antineumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200183525-A2.

XX XX 08-NOV-2001.

XX PF 02-MAY-2001; 2001WO-US014310.

XX PR 03-MAY-2000; 2000US-00563286.

XX PA (AMGEN-) AMGEN INC.

XX PI Peige U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX DR MPI; 2002-130313/17.

XX PT Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility.

XX PS Claim 39; Page 44; 176pp; English.

XX CC The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antineumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 81; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWLTTC 14
 |||||
 Db 1 CELVGPMSLMSWLTTC 14

```

RESULT 7
ADJ73048
ID ADJ73048 standard; peptide; 14 AA.
XX AC
XX ADJ73048;
XX AC
DT 06-MAY-2004 (first entry)
XX DE
XX TPO mimetic peptide sequence SeqID 502.
XX XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
XX KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
XX KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
XX KW TPO.
XX OS Synthetic.
XX OS
XX WO2003084477-A2.
XX PN
XX 16-OCT-2003.
XX PD
XX 24-MAR-2003; 2003WO-US009139.
XX PF
XX 29-MAR-2002; 2002US-0368791P.
XX PR
XX (CENZ ) CENTOCOR INC.
XX PA
XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX PI WPI; 2003-804237/75.
XX DR
XX New CDR mimetibody comprising a portion of a heavy or light chain
XX PT variable region comprising human framework or ligand binding region,
XX PT useful for preparing a composition for treating e.g., immune,
XX PT cardiovascular or neurologic disease.
XX PS Disclosure; SEQ ID NO 502; 97pp; English.
XX XX
XX This invention relates to novel mammalian CDR mimetibodies, specific
XX CC portions or variants thereof. Specifically, it refers to an antibody
XX CC fragment where a protein has been inserted into, or replaces a portion
XX CC of, one or more CDR regions, such that each CDR mimetibody comprises at
XX CC least one portion of a heavy chain or light chain variable region, which
XX CC itself comprises at least one human framework region and at least one
XX CC ligand binding region (LBR). The present invention describes human
XX CC mimetibodies, including modified immunoglobulins and cleavage products
XX CC that can be useful in gene therapy and the generation of transgenic
XX CC plants and animals. Furthermore, the CDR mimetibody is useful for
XX CC preparing compositions for modulating, treating or reducing the symptoms
XX CC of immune, cardiovascular, infectious, malignant and/or neurologic
XX CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
XX CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
XX CC peptide sequence is a TPO mimetic peptide sequence used to make a
XX CC mimetibody of the invention.
XX SQ Sequence 14 AA;
XX Query Match 100.0%; Score 81; DB 7; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-06;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CELVGPSSLMSWLTTC 14
DB 1 CELVGPSSLMSWLTTC 14
| | | | | | | | | |
| | | | | | | | | |
RESULT 8
ADJ52683
ID ADJ52683 standard; peptide; 14 AA.
XX XX
XX ADJ52683;
XX AC
XX CH1 deleted mimetibody-related peptide SeqID502.
DT 06-MAY-2004 (first entry)
XX DE
XX CH1 deleted mimetibody-related peptide SeqID502.
XX KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
XX KW fungicide; gene therapy; immune disorder; cardiovascular disease;
XX KW arrhythmia; hypertension; heart failure; neurodegenerative;
XX KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
XX KW cancerous condition; infectious disease; bacterial infection;
XX KW viral infection; fungal infection.
XX OS Unidentified.
XX OS Synthetic.
XX WO2004002417-A2.
XX PN
XX 08-JAN-2004.
XX PD
XX 27-JUN-2003; 2003WO-US020347.
XX PF
XX 28-JUN-2002; 2002US-0392431P.
XX PR
XX (CENZ ) CENTOCOR INC.
XX PA
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX PI Kutoloski KA;
XX PI WPI; 2004-082870/08.
XX DR
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
XX PT modulating, treating, alleviating, preventing an immune, cardiovascular,
XX PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
XX PT diseases.
XX PS Claim 2; SEQ ID NO 502; 129pp; English.
XX XX
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
XX CC which encode them), compositions, methods and uses. The invention may be
XX CC useful for the development of compounds with an immunosuppressive,
XX CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
XX CC antibacterial, virucide or fungicide activity. In addition, the disclosed
XX CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
XX CC is useful for diagnosing or treating a disease condition in a cell,
XX CC tissue, organ or animal, specifically for modulating, treating,
XX CC alleviating, preventing the incidence or reducing the symptoms of an
XX CC immune, cardiovascular (for example arrhythmia, hypertension or heart
XX CC failure), or neurodegenerative (for example multiple sclerosis, dementia
XX CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
XX CC conditions, or infectious diseases (for example bacterial, viral or
XX CC fungal infection). The present sequence is that of a peptide which may be
XX CC used during the creation of a mimetibody of the invention.
XX SQ Sequence 14 AA;
XX Query Match 100.0%; Score 81; DB 8; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-06;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CELVGPSSLMSWLTTC 14
DB 1 CELVGPSSLMSWLTTC 14
| | | | | | | | | |
| | | | | | | | | |
RESULT 9
ADJ51644
ID ADJ51644 standard; peptide; 14 AA.
XX XX
XX ADJ51644;
XX AC
XX 06-MAY-2004 (first entry)
XX DT
XX CH1 deleted mimetibody-related peptide SeqID502.
XX DE

```


Query Match 65.4%; Score 53; DB 7; Length 25;
Best Local Similarity 57.1%; Pred. No. 0.15;

	SQ	Sequence 14 AA;	
	Query Match	64.2%; Score 52; DB 5; Length 14;	
	Best Local Similarity	57.1%; Pred. No. 0.12;	
	Matches	8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;	
QY	1	CELVGPSLMSWLTG 14	
DB	1	CEQDGTLLLEWLKC 14	
	RESULT 14		
ID	ADJ73047		
XX	ADJ73047 standard; peptide; 14 AA.		
AC	ADJ73047;		
XX	06-MAY-2004 (first entry)		
DT	TPO mimetic peptide sequence SeqID 501.		
XX	mimetic; CDR mimetibody; gene therapy; transgenic; immune;		
XX	cardiovascular; infectious; malignant; neurologic disease; anaemia;		
KW	immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;		
KW	TPO.		
OS	Synthetic.		
XX	WO2003084477-A2.		
PN	16-OCT-2003.		
PD	24-MAR-2003; 2003WO-US009139.		
PX	29-MAR-2002; 2002US-0368791P.		
PR	(CENZ) CENTOCOR INC.		
XX	Heavner GA, Knight DM, Scallion BJ, Ghayeb J;		
XX	WPI; 2003-804237/75.		
DR	New CDR mimetibody comprising a portion of a heavy or light chain		
PT	variable region comprising human framework or ligand binding region,		
PT	useful for preparing a composition for treating e.g., immune,		
PT	cardiovascular or neurologic disease.		
XX	Disclosure; SEQ ID NO 501; 97pp; English.		
PS	This invention relates to novel mammalian CDR mimetibodies, specific		
XX	portions or variants thereof. Specifically, it refers to an antibody		
CC	fragment where a protein has been inserted into, or replaces a portion		
CC	of, one or more CDR regions, such that each CDR mimetibody comprises at		
CC	least one portion of a heavy chain or light chain variable region, which		
CC	itself comprises at least one human framework region and at least one		
CC	ligand binding region (LBR). The present invention describes human		
CC	mimetibodies, including modified immunoglobulins and cleavage products		
CC	that can be useful in gene therapy and the generation of transgenic		
CC	plants and animals. Furthermore, the CDR mimetibody is useful for		
CC	preparing compositions for modulating, treating or reducing the symptoms		
CC	of immune, cardiovascular, infectious, malignant and/or neurologic		
CC	diseases, as well as anaemia. Accordingly, they exhibit immunomodulatory,		
CC	cardiant, antimicrobial, cytostatic and neuroprotective activities. This		
CC	peptide sequence is a TPO mimetic peptide sequence used to make a		
CC	mimetibody of the invention.		
XX	Sequence 14 AA;		
	Query Match	64.2%; Score 52; DB 7; Length 14;	
	Best Local Similarity	57.1%; Pred. No. 0.12;	
	Matches	8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;	
QY	1	CELVGPSLMSWLTG 14	

```

Db      1  CEQDGFLLLEWLKC 14
|| ||:|: || |
|| ||:|: || |

RESULT 15
ADJ52682
ID ADJ52682 standard; peptide; 14 AA.
XX
AC ADJ52682;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID501.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghayeb J, Scallan BJ, Nesspor TC;
PI Kutolowski KA;
XX
PI WPI; 2004-082870/08.
XX
DR New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 2; SEQ ID NO 501; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 14 AA;

Query Match      64.2%; Score 52; DB 8; Length 14;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 8; Conservativity 2; Mismatches 4; Indels 0; Gaps 0;

QY      1  CELVGPGLMSWLTC 14
|| ||:|: || |
|| ||:|: || |

Db      1  CEQDGFLLLEWLKC 14

```

RESULT 16

ADJ51643

ID ADJ51643 standard; peptide; 14 AA.

XX

AC ADJ51643;

XX

DT 06-MAY-2004 (first entry)

XX

DE CH1 deleted mimetibody-related peptide SeqID501.

XX

KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;

KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;

KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;

KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;

KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;

KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;

KW dental disorder; oral disorder; dermatological disorder; ear disorder;

KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;

KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;

KW obstetric disorder; haematologic disorder; immunologic disorder;

KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;

KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;

KW renal disorder; pulmonary disorder.

XX

OS Unidentified.

OS Synthetic.

XX

PN WO2004002424-A2.

XX

PD 08-JAN-2004.

XX

PF 30-JUN-2003; 2003WO-US020495.

XX

PR 28-JUN-2002; 2002US-0392431P.

XX

PR 19-SEP-2002; 2002US-0412144P.

XX

XX (CENZ) CENTOCOR INC.

XX

XX Heavner GA, Knight DM, Ghayeb J, Scallan BJ, Nesspor TC;

PI Kutolowski KA;

XX

DR WPI; 2004-082872/08.

XX

XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for

PT diagnosing, preventing or treating cardiovascular, dermatologic,

PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and

PT nutritional disorders.

XX

PS Claim 15; SEQ ID NO 501; 123pp; English.

XX

CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences

CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,

CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,

CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,

CC immunomodulatory, antiallergic, muscular-Gen, cytostatic,

CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or

CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

CC modulator or cytokine-agonist. The methods and compositions of the

CC present invention are useful for the diagnosis, prevention and/or

CC treatment of diseases or conditions associated with aberrant expression

CC or activity of the CH1 deleted mimetibody, such as a bone or joint,

CC cardiovascular, dental or oral, dermatological, ear, nose or throat,

CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,

CC obstetric, haematologic, immunologic, allergic, infectious,

CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC pediatric, psychiatric, renal or pulmonary disorders. The present

CC sequence is that of a peptide which may be used during the creation of a

CC mimetibody of the invention.

XX

SQ Sequence 14 AA;

Query Match 64.2%; Score 52; DB 8; Length 14;
 Best Local Similarity 57.1%; Pred. No. 0.12; Mismatches 2; Gaps 0;
 Matches 8; Conservative 2; Indels 4; Indels 0; Gaps 0;

QY 1 CELVGPLMSWLTC 14
 |||||
 DB 1 CEQDGPILLEWLKC 14
 |||||

RESULT 17
 ABB09517
 ID ABB09517 standard; protein; 129 AA.
 XX
 AC ABB09517;
 XX
 DT 01-NOV-2002 (first entry)
 XX
 DE Human RNase-like NOV8 protein, SEQ ID NO:32.
 XX
 KW Human; NOVX; neurological disorder; Alzheimer's disease;
 KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;
 KW addition; tuberous sclerosis; cancer; immune disorder; allergy;
 KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;
 KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;
 KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;
 KW pancreatitis; cirrhosis; glomerular endotheiosis; bacterial infection;
 KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;
 KW metabolic pathway regulation disorder; cytostatic; neuroprotective;
 KW antiinflammatory; immunosuppressive; analgesic; antiatherosclerotic;
 KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;
 KW differentiation; proliferation; motility; haematopoiesis; wound healing;
 KW angiogenesis; forensic biology; transgenic animal; drug screening;
 KW gene therapy; NOV8; RNase-like; chromosome 14.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT Misc-difference 28
 FT /note= "This residue is Arg in a polymorphic variant"
 FT
 FT Misc-difference 64
 FT /note= "This residue is Thr in a polymorphic variant"
 FT
 FT Misc-difference 69
 FT /note= "This residue is Leu in a polymorphic variant"
 FT
 XX W020253742-A2.
 XX
 XX 11-JUL-2002.
 XX
 XX 07-JAN-2002; 2002WO-US000375.
 XX
 XX 05-JAN-2001; 2001US-0260018P.
 XX 08-JAN-2001; 2001US-0260360P.
 XX 28-FEB-2001; 2001US-0272411P.
 XX 02-MAR-2001; 2001US-0272817P.
 XX 05-JUL-2001; 2001US-0303231P.
 XX 12-JUL-2001; 2001US-0305060P.
 XX 10-SEP-2001; 2001US-0318405P.
 XX 12-SEP-2001; 2001US-0318700P.
 XX 04-JAN-2002; 2002US-00037417.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Kekuda R, Alsobrook JP, Tchernev VT, Liu X, Spytek KA;
 XX Patturajan M, Grosse WM, Lepley DM, Burgess CE, Vernet CAM, Li L;
 XX Gorman L, Edinger S, Schiore P, Ellerman K, Malyankar U;
 XX Rothenberg M, Stone D, Boldog F, Guo X, Shenoy S, Anderson D;
 XX Padigar M, Taupier RU, Miller CE, Eisen A;
 XX
 XX WPI; 2002-583619/62.
 XX N-PSDB; ABQ93895.
 XX

PT Novel polypeptides and nucleic acids homologous to transmembrane
 PT receptor, thymosin, neuromodulin-like family of proteins for diagnosing,
 PT treating cancer, atherosclerosis, neurological, skin and autoimmune
 PT disorders.
 XX
 XX Claim 1c; Page 99; 323pp; English.
 XX
 XX The invention relates to 24 novel human proteins designated NOV1-NOV14
 CC (ABB09501-ABB09524), collectively referred to as NOVX proteins, and
 CC nucleic acids encoding them (ABQ93879-ABQ93902). NOVX proteins, and
 CC nucleotides are useful in the treatment, diagnosis or prevention of NOVX-
 CC associated disorders or in the manufacture of a medicament for treating
 CC such disorders, with specific applications described for each of the 24
 CC NOVX proteins, based on their homology to known proteins. Various
 CC disorders are associated with NOVX proteins including neurological
 CC disorders (e.g., Alzheimer's, Huntington's and Parkinson's diseases),
 CC pain, behavioural disorders, addiction, tuberous sclerosis, cancers
 CC (e.g., colorectal cancer, leukaemia and osteosarcoma), immune disorders
 CC (e.g., allergies and autoimmune diseases), myasthenia gravis, asthma,
 CC various forms of arthritis, diabetes, thyroiditis, cardiovascular disease
 CC (e.g., hypertension), reproductive disorders, endometriosis,
 CC incontinence, psoriasis, scleroderma, alopecia, ulcers, pancreatitis,
 CC cirrhosis, glomerular endotheiosis, polycystic kidney disease, endocrine
 CC disorders, obesity, bacterial infections and particularly cardiomyopathy,
 CC atherosclerosis, cell signal processing-related disorders and disorders
 CC of metabolic pathway regulation. NOVX nucleic acids and polypeptides may
 CC be used to identify cellular receptors or downstream effectors which
 CC binds to a NOVX protein, and are also useful as targets for the
 CC identification of small molecules that modulate or inhibit processes such
 CC as neurogenesis, cell differentiation, cell motility, cellular
 CC proliferation, haematopoiesis, wound healing and angiogenesis. NOVX
 CC nucleic acid sequences can be used to identify a cell or tissue type and
 CC are useful as a source of primers or probes for forensic biology and for
 CC identifying and cloning NOVX homologues in other cell types. Cells
 CC comprising NOVX nucleic acids are useful for producing non-human
 CC transgenic animals which are useful for studying the function and
 CC activity of NOVX proteins and for identifying and evaluating modulators
 CC of NOVX activity. The present sequence represents the RNase-like protein
 CC NOV8. The gene encoding NOV8 is located on chromosome 14
 XX
 SQ Sequence 129 AA;
 Query Match 60.5%; Score 49; DB 5; Length 129;
 Best Local Similarity 75.0%; Pred. No. 3.9;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPLMSWLTC 14
 |||||
 DB 114 LTGHSLSMWLVC 125
 |||||

RESULT 18
 AAM29099
 ID AAM29099 standard; protein; 198 AA.
 XX
 AC AAM29099;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #3136 encoded by probe for measuring placental gene expression.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200157272-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US000663.
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 XX

		The invention relates to novel human APP proteins (not defined) and the nucleic acids that encode them. APP proteins are useful as standards in assays of protein and protein inhibitors in both clinical and research settings, as protein and amino acid supplements, including hydrolysates. The nucleic acids are useful for radiation hybrid mapping. Secretary fusion proteins of APP are useful in cancer therapy, for enhancing in vitro cytotoxicity, for enhancing in vivo killing of target tissues, and for targeted cells or tissue inhibition or ablation. Anti-APP antibodies are useful for isolating target polypeptides by affinity purification, in diagnostic assays for determining circulating or localised levels of target polypeptides, for tissue typing, for cell sorting, for screening expression libraries, for generating anti-idiotypic antibodies, and as neutralising antibodies or as antagonists to block protein activity in vitro and in vivo. APP proteins and nucleic acids may be used to diagnose or treat (e.g. by gene therapy) diseases associated with the malfunction of the APP e.g. AβP6924 and Marfan's syndrome, AβP576853/AβP39158 and CC thrombocytopaenia, leukaemia, porphyria, Gilles De La Tourette's CC syndrome, AβP652829 and squamous cell carcinoma, diabetes mellitus, CC Grave's disease, AβP664311 and colon cancer, AβP253034 and Alzheimer's CC disease and AβP686580 and epiphyseal dysplasia. Many more examples of CC disease are given in the specification. The present sequence represents CC an APP of the invention
XX		
SQ	Sequence 199 AA;	
	Query Match 60.5%; Score 49; DB 4; Length 199;	
	Best Local Similarity 75.0%; Pred. No. 6.2;	
	Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0	
QY	3 LVGPSLMSWLTCTC 14	
Dd	 184 LTGHSLSWLVC 195	
RESULT 20		
AAG77853		
ID	AAG77853 standard; protein; 199 AA.	
XX		
AC	AAG77853;	
XX		
DT	04-JAN-2002 (first entry)	
XX		
DE	Human ribonuclease Zrnase1.	
XX		
KW	Ribonuclease; human; Zrnase1; gene therapy; tumour; inflammation; gout; anemia; Addison's disease; allergy; asthma; arteriosclerosis; bronchitis; Crohn's disease; ulcerative colitis; atopic dermatitis; Grave's disease; diabetes mellitus; glomerulonephritis; irritable bowel syndrome; lupus erythematosus; multiple sclerosis; myasthenia gravis; myocardial infarction; pericardial inflammation; osteoarthritis; osteoporosis; pancreatitis; rheumatoid arthritis; scleroderma; Sjogren's syndrome; autoimmune thyroiditis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	Peptide 1..19	
FT	Protein /label= signal_sequence	
FT	20..199	
FT	/label= mature_Zrnase1	
XX		
PN	WO200166710-A2.	
XX		
PD	13-SEP-2001.	
XX		
PF	07-MAR-2001; 2001WO-US007333.	
XX		
PR	08-MAR-2000; 2000US-00522765.	
XX	(ZYMO) ZYMOGENETICS INC.	
PA		
XX		
PI	Conklin DC;	

XX WPI: 2001-596833/67.
 DR N-PSDB; AAH78831.
 XX Novel isolated Zrnasei polypeptide which is a member of the human
 PT ribonuclease family, useful for treating tumors, Addison's disease,
 PT Crohn's disease, Grave's disease, Sjogren's syndrome and
 PT glomerulonephritis.
 XX Claim 3; Page 2; 89pp; English.
 XX The sequence represents the novel human ribonuclease Zrnasei protein. The
 CC invention relates to a novel isolated Zrnasei polypeptide which is a
 CC member of the human ribonuclease family, having a use in gene therapy.
 CC The Zrnasei polypeptide, nucleotides, antibodies and fusion proteins of
 CC the invention are useful for preventing or treating a disorder associated
 CC with excessive cellular proliferation such as tumour, and to prevent or
 CC to treat inflammation associated disorders such as Addison's disease,
 CC allergies, anaemia, asthma, atherosclerosis, bronchitis, Crohn's disease,
 CC ulcerative colitis, atopic dermatitis, diabetes mellitus,
 CC glomerulonephritis, gout, Grave's disease, irritable bowel syndrome,
 CC lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial
 CC or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
 CC rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune
 CC thyroiditis. The polypeptide and polynucleotide may also be used as
 CC educational tools in laboratory practicum kits for courses related to
 CC genetics and molecular biology, protein chemistry and antibody production
 CC and analysis
 XX SQ Sequence 199 AA;
 Query Match 60.5%; Score 49; DB 4; Length 199;
 Best Local Similarity 75.0%; Pred. No. 6.2;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 LVGPSLSWLTC 14
 DB 184 LTGHSLSWLVC 195
 RESULT 21
 AAU29085
 ID AAU29085 standard; protein; 199 AA.
 XX AAU29085;
 XX 18-DEC-2001 (first entry)
 XX Human PRO polypeptide sequence #62.
 XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
 XX Homo sapiens.
 XX WO200168848-A2.
 XX 20-SEP-2001.
 XX 28-FEB-2001; 2001WO-US0006520.
 XX 01-MAR-2000; 2000WO-US0005601.
 PR 02-MAR-2000; 2000WO-US0005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 06-MAR-2000; 2000US-0186968P.
 PR 14-MAR-2000; 2000US-0189320P.
 PR 14-MAR-2000; 2000US-0189328P.
 PR 15-MAR-2000; 2000WO-US0006884.
 PR 21-MAR-2000; 2000US-0190828P.
 PR 21-MAR-2000; 2000US-0191007P.
 PR 21-MAR-2000; 2000US-0191048P.

PR 21-MAR-2000; 2000US-0191314P.
 PR 28-MAR-2000; 2000US-0192855P.
 PR 29-MAR-2000; 2000US-0193032P.
 PR 30-MAR-2000; 2000WO-US0008439.
 PR 04-APR-2000; 2000US-0194449P.
 PR 04-APR-2000; 2000US-0194647P.
 PR 11-APR-2000; 2000US-0195975P.
 PR 11-APR-2000; 2000US-0196000P.
 PR 11-APR-2000; 2000US-0196187P.
 PR 11-APR-2000; 2000US-0196690P.
 PR 11-APR-2000; 2000US-0196820P.
 PR 18-APR-2000; 2000US-0198121P.
 PR 18-APR-2000; 2000US-0198585P.
 PR 25-APR-2000; 2000US-0199397P.
 PR 25-APR-2000; 2000US-0199550P.
 PR 25-APR-2000; 2000US-0199654P.
 PR 03-MAY-2000; 2000US-0201516P.
 PR 17-MAY-2000; 2000WO-US011705.
 PR 22-MAY-2000; 2000WO-US0114042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 22-AUG-2000; 2000US-00644848.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI: 2001-602746/68.
 DR N-PSDB; AAS45986.
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumors, such as prostate and breast tumors, in mammals and to
 PT screen for modulators of the compounds.
 XX Claim 11; Fig 124; 774pp; English.
 XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders
 XX SQ Sequence 199 AA;
 Query Match 60.5%; Score 49; DB 4; Length 199;
 Best Local Similarity 75.0%; Pred. No. 6.2;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 LVGPSLSWLTC 14
 DB 184 LTGHSLSWLVC 195
 RESULT 22

AAG89149
ID AAG89149 standard; protein; 199 AA.
XX
AC AAG89149;
XX
XX 11-SEP-2001 (first entry)
XX
XX Human secreted protein, SEQ ID NO: 269.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET.
XX
XX Homo sapiens.
XX
XX WO200142451-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-IB001938.
XX
XX 08-DEC-1999; 99US-0169629P.
PR
XX 06-MAR-2000; 2000US-0187470P.
PR
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI
XX WPI; 2001-367870/38.
XX
XX N-PSDB; AAH64752.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
XX Claim 21; Page 808; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased GENSET
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patient's own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET polypeptide of
CC the invention
XX
XX Sequence 199 AA;
SQ
Query Match 60.5%; Score 49; DB 4; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPMSLMSWLTLC 14
Db 184 LTGHSLMSWLVC 195
RESULT 23
ABP43532
ID ABP43532 standard; protein; 199 AA.
XX
AC ABP43532;
XX
XX 08-AUG-2002 (first entry)
DT

XX Human secreted protein (SECP) 56.
DE
XX
XX Human; secreted protein; SECP; SECP expression; gene therapy;
KW protein therapy; immune system disorders; AIDS; thymic hypoplasia;
KW anaemia; asthma; Crohn's disease; neurological disorder; epilepsy;
KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
KW developmental disorder; cell proliferative disorder; cancer.
XX
XX Homo sapiens.
XX
XX WO200226982-A2.
XX
XX 04-APR-2002.
XX
XX 25-SEP-2001; 2001WO-US030042.
XX
XX 29-SEP-2000; 2000US-0236869P.
PR
XX 11-OCT-2000; 2000US-0239812P.
PR
XX 12-OCT-2000; 2000US-0240108P.
PR
XX 17-OCT-2000; 2000US-0241282P.
PR
XX 20-OCT-2000; 2000US-0242218P.
PR
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;
PI Sanjanwala MS, Walia NK, Baughn MR, Sapperstein SK, Lal P;
PI Thornton M, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;
PI Thangavelu K, Gietzen KJ, Ding L, Au-Young J, Tran B, Policky JL;
PI Lee S, Lu DAM, Burford NJ, Warren BA, Gururajan R, Duggan BM;
PI Honchell CD, Hafalia AJA;
XX
XX WPI: 2002-394239/42.
XX
XX N-PSDB; ABN99415.
XX
XX New human secreted proteins, useful for diagnosing, treating or
PT preventing immune system disorders (e.g. Crohn's disease), neurological
PT disorders (e.g. Parkinson's disease), or cell proliferative disorders
PT (e.g. cancers).
XX
XX Claim 1; Page 187-188; 238pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the
CC invention are useful for treating/preventing disorders associated with
CC decreased or elevated expression of SECP. The SECP DNA and protein
CC sequences are specifically useful for treating/preventing (i.e. gene
CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic
CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders
CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);
CC developmental disorders (e.g. Down's syndrome); and cell proliferative
CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the
CC human secreted proteins (SECP) of the invention
XX
XX Sequence 199 AA;
SQ
Query Match 60.5%; Score 49; DB 5; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPMSLMSWLTLC 14
Db 184 LTGHSLMSWLVC 195
RESULT 24
AAU12076
ID AAU12076 standard; protein; 199 AA.
XX
AC AAU12076;
XX
XX 23-APR-2002 (first entry)
DT

DE Human ribonuclease Zrnase1 polypeptide.
XX
XX Human, ribonuclease; Zrnase1; cellular proliferation; tumour; cancer;
KW inflammatory disorder; allergy; asthma; atherosclerosis; cytostatic;
KW autoimmune disorder; antiinflammatory; cardiant; enzyme.
XX
OS Homo sapiens.
XX
XX US2001049434-A1.
XX
XX 06-DEC-2001.
XX
XX 07-MAR-2001; 2001US-00801231.
XX
XX 08-MAR-2000; 2000US-0187917P.
PR (CONK/) CONKLIN D C.
XX Conklin DC;
XX
XX WPI; 2002-121402/16.
DR N-PSDB; AAS15496.
XX
XX Ribonuclease, Zrnase1, useful for preventing and treating, e.g., cancer,
PT diabetes, asthma and osteoarthritis, and to identify and to isolate
PT Zrnase1 substrates and inhibitors.
XX
XX Claim 3; Page 1; 37pp; English.
XX
XX The present invention relates to the isolation of a novel ribonuclease,
CC Zrnase1. Zrnase1 and fusion proteins comprising it can be used to prevent
CC or to treat a disorder associated with excessive cellular proliferation.
CC Examples of such disorders include various types of tumours (e.g.
CC adenocarcinoma, leukaemia, lymphoma and melanoma). In addition, Zrnase1
CC can be used to prevent or to treat inflammation associated disorders,
CC such as Addison's disease, adult respiratory distress syndrome (ARDS),
CC allergies, anaemia, asthma, atherosclerosis, bronchitis, cholecystitis,
CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
CC diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis,
CC gout, Graves' disease, hypersinophilia, irritable bowel syndrome, lupus
CC erythematosus, multiple sclerosis, myasthenia gravis, myocardial or
CC pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
CC polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome and
CC autoimmune thyroiditis. Zrnase1 can also be used to identify and to
CC isolate Zrnase1 substrates and inhibitors. The present sequence
CC represents human ribonuclease, Zrnase1
XX
SQ Sequence 199 AA;

Query Match 60.5%; Score 49; DB 5; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSTLSWLTTC 14
Db 184 LTGHSLSWLVLC 195

RESULT 25
ABUS8693
ID ABUS8693 standard; protein; 199 AA.
XX
XX AC ABUS8693;
XX
XX 15-APR-2003 (first entry)
XX
XX Human PRO polypeptide #294.
XX
XX Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
XX Homo sapiens.
OS

XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083495P.
PR 23-APR-1998; 98US-0083496P.
PR 23-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084840P.
PR 07-MAY-1998; 98US-0084843P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.

```
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 05-JUN-1998; 98US-0088655P.
PR 05-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088862P.
PR 11-JUN-1998; 98US-0088867P.
PR 12-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090533P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091488P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
```

```
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 25-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
```

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
| | | | | | | |
Db 184 LTGHSLMSWLVC 195

RESULT 26

ABU88241
ID ABU88241 standard; protein; 199 AA.
XX
AC ABU88241;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO19669.
XX
KW Human; secreted and transmembrane protein; PRO; Gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release; tumour;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX

PF 26-JUN-2002; 2002US-00183012.
XX 18-SEP-1997; 97US-0059463P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069435P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0077888P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083435P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 15-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085582P.
PR 18-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087088P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087837P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 15-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090890P.
PR 25-JUN-1998; 98US-0090894P.
PR 25-JUN-1998; 98US-0090959P.
PR 25-JUN-1998; 98US-0090965P.
PR 25-JUN-1998; 98US-0090966P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.

PR	10-SEP-1998;	98US-0099763P.	PR	28-OCT-1997;	97US-0063541P.
PR	10-SEP-1998;	98US-0099812P.	PR	28-OCT-1997;	97US-0063544P.
PR	15-SEP-1998;	98US-0100388P.	PR	28-OCT-1997;	97US-0063564P.
PR	16-SEP-1998;	98US-0100662P.	PR	29-OCT-1997;	97US-0063734P.
PR	16-SEP-1998;	98US-0100664P.	PR	31-OCT-1997;	97US-0063870P.
PR	16-SEP-1998;	98US-0101751P.	PR	31-OCT-1997;	97US-0064103P.
PR	16-SEP-1998;	98MO-US019130.	PR	13-NOV-1997;	97US-0065311P.
PR	17-SEP-1998;	98US-0100683P.	PR	21-NOV-1997;	97US-0066120P.
PR	17-SEP-1998;	98US-0100684P.	PR	24-NOV-1997;	97US-0066466P.
PR	17-SEP-1998;	98US-0100919P.	PR	24-NOV-1997;	97US-0066772P.
PR	17-SEP-1998;	98US-0100930P.	PR	11-DEC-1997;	97US-0069335P.
PR	17-SEP-1998;	98US-0100930P.	PR	12-DEC-1997;	97US-0069425P.
PR	18-SEP-1998;	98US-0100849P.	PR	17-DEC-1997;	97US-0069870P.
PR	18-SEP-1998;	98US-0101014P.	PR	18-DEC-1997;	97US-0068017P.
PR	18-SEP-1998;	98US-0101058P.	PR	10-MAR-1998;	98US-0077450P.
PR	23-SEP-1998;	98US-0101471P.	PR	11-MAR-1998;	98US-0077632P.
PR	23-SEP-1998;	98US-0101472P.	PR	11-MAR-1998;	98US-0077649P.
PR	23-SEP-1998;	98US-0101475P.	PR	20-MAR-1998;	98US-0078886P.
PR	23-SEP-1998;	98US-0101477P.	PR	20-MAR-1998;	98US-0078939P.
PR	24-SEP-1998;	98US-0101738P.	PR	27-MAR-1998;	98US-0079664P.
PR	24-SEP-1998;	98US-0101739P.	PR	27-MAR-1998;	98US-0079786P.
PR	24-SEP-1998;	98US-0101743P.	PR	31-MAR-1998;	98US-0080107P.
PR	24-SEP-1998;	98US-0101922P.	PR	31-MAR-1998;	98US-0080194P.
PR	25-SEP-1998;	98US-0101786P.	PR	01-APR-1998;	98US-0080327P.
PR	29-SEP-1998;	98US-0102207P.	PR	01-APR-1998;	98US-0080333P.
PR	29-SEP-1998;	98US-0102240P.	PR	08-APR-1998;	98US-0081049P.
PR	29-SEP-1998;	98US-0102330P.	PR	08-APR-1998;	98US-0081070P.
PR	29-SEP-1998;	98US-0102331P.	PR	09-APR-1998;	98US-0081195P.
PR	29-SEP-1998;	98US-0102487P.	PR	15-APR-1998;	98US-0081838P.
PR	30-SEP-1998;	98US-0102570P.	PR	15-APR-1998;	98US-0082568P.
PR	30-SEP-1998;	98US-0102571P.	PR	21-APR-1998;	98US-0082569P.
PR	01-OCT-1998;	98US-0102684P.	PR	22-APR-1998;	98US-0082704P.
PR	01-OCT-1998;	98US-0102684P.	PR	22-APR-1998;	98US-0082797P.
PR	02-OCT-1998;	98US-0102687P.	PR	28-APR-1998;	98US-0083322P.
PR	06-OCT-1998;	98US-0102965P.	PR	29-APR-1998;	98US-0083495P.
PR	06-OCT-1998;	98US-0103258P.	PR	29-APR-1998;	98US-0083496P.
PR	06-OCT-1998;	98US-0103259P.	PR	29-APR-1998;	98US-0083499P.
PR	06-OCT-1998;	98US-0103259P.	PR	29-APR-1998;	98US-0083559P.
PR	06-OCT-1998;	98US-0103259P.	PR	05-MAY-1998;	98US-0084366P.
PR	06-OCT-1998;	98US-0103259P.	PR	06-MAY-1998;	98US-0084414P.
PR	06-OCT-1998;	98US-0103259P.	PR	07-MAY-1998;	98US-0084639P.
PR	06-OCT-1998;	98US-0103259P.	PR	07-MAY-1998;	98US-0084640P.
PR	06-OCT-1998;	98US-0103259P.	PR	07-MAY-1998;	98US-0084643P.
PR	06-OCT-1998;	98US-0103259P.	PR	15-MAY-1998;	98US-0085579P.
PR	06-OCT-1998;	98US-0103259P.	PR	15-MAY-1998;	98US-0085580P.
PR	06-OCT-1998;	98US-0103259P.	PR	15-MAY-1998;	98US-0085582P.
PR	06-OCT-1998;	98US-0103259P.	PR	15-MAY-1998;	98US-0085700P.
PR	06-OCT-1998;	98US-0103259P.	PR	18-MAY-1998;	98US-0086023P.
PR	06-OCT-1998;	98US-0103259P.	PR	22-MAY-1998;	98US-0086392P.
PR	06-OCT-1998;	98US-0103259P.	PR	22-MAY-1998;	98US-0086486P.
PR	06-OCT-1998;	98US-0103259P.	PR	28-MAY-1998;	98US-0087098P.
PR	06-OCT-1998;	98US-0103259P.	PR	28-MAY-1998;	98US-0087208P.
PR	06-OCT-1998;	98US-0103259P.	PR	02-JUN-1998;	98US-0087609P.
PR	06-OCT-1998;	98US-0103259P.	PR	02-JUN-1998;	98US-0087759P.
PR	06-OCT-1998;	98US-0103259P.	PR	03-JUN-1998;	98US-0087827P.
PR	06-OCT-1998;	98US-0103259P.	PR	04-JUN-1998;	98US-0088025P.
PR	06-OCT-1998;	98US-0103259P.	PR	04-JUN-1998;	98US-0088028P.
PR	06-OCT-1998;	98US-0103259P.	PR	04-JUN-1998;	98US-0088029P.
PR	06-OCT-1998;	98US-0103259P.	PR	04-JUN-1998;	98US-0088033P.
PR	06-OCT-1998;	98US-0103259P.	PR	04-JUN-1998;	98US-0088326P.
PR	06-OCT-1998;	98US-0103259P.	PR	05-JUN-1998;	98US-0088167P.
PR	06-OCT-1998;	98US-0103259P.	PR	05-JUN-1998;	98US-0088202P.
PR	06-OCT-1998;	98US-0103259P.	PR	05-JUN-1998;	98US-0088212P.
PR	06-OCT-1998;	98US-0103259P.	PR	05-JUN-1998;	98US-0088217P.
PR	06-OCT-1998;	98US-0103259P.	PR	09-JUN-1998;	98US-0088655P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088722P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088738P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088740P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088811P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088824P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088825P.
PR	06-OCT-1998;	98US-0103259P.	PR	10-JUN-1998;	98US-0088826P.
PR	06-OCT-1998;	98US-0103259P.	PR	11-JUN-1998;	98US-0088861P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLSWLTLC 14
DB 184 LTGHSLMSWLC 195

RESULT 27
ABU84556
ID ABU84556 standard; protein; 199 AA.
XX AC ABU84556;
XX DT 02-AUG-2003 (first entry)
XX DE Human secreted/transmembrane protein (PRO) #294.
XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX OS Homo sapiens.
XX PN US2003032112-A1.
XX PD 13-FEB-2003.
XX PF 21-JUN-2002; 2002US-00176756.
XX PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.

```
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089558P.
PR 17-JUN-1998; 98US-0089553P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100682P.
PR 16-SEP-1998; 98US-0100684P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089558P.
PR 17-JUN-1998; 98US-0089553P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100682P.
PR 16-SEP-1998; 98US-0100684P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLSMLTC 14
Db 184 LTGHSLSMLVC 195
RESULT 28
ABR66430
ID ABR66430 standard; protein; 199 AA.
XX
AC ABR66430;
XX
DT 05-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027278-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176987.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
```


PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084633P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085582P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091544P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.

PR	18-SEP-1998;	98US-0101014P.	PR	11-DEC-1997;	97US-0069335P.
PR	18-SEP-1998;	98US-0101069P.	PR	12-DEC-1997;	97US-0069425P.
PR	23-SEP-1998;	98US-0101471P.	PR	17-DEC-1997;	97US-0069870P.
PR	23-SEP-1998;	98US-0101472P.	PR	18-DEC-1997;	97US-0068017P.
PR	23-SEP-1998;	98US-0101475P.	PR	10-MAR-1998;	98US-0077450P.
PR	23-SEP-1998;	98US-0101477P.	PR	11-MAR-1998;	98US-0077632P.
PR	24-SEP-1998;	98US-0101738P.	PR	11-MAR-1998;	98US-0077649P.
PR	24-SEP-1998;	98US-0101739P.	PR	20-MAR-1998;	98US-0078886P.
PR	24-SEP-1998;	98US-0101743P.	PR	20-MAR-1998;	98US-0078939P.
PR	24-SEP-1998;	98US-0101922P.	PR	27-MAR-1998;	98US-0079664P.
PR	25-SEP-1998;	98US-0101786P.	PR	27-MAR-1998;	98US-0079786P.
PR	29-SEP-1998;	98US-0102240P.	PR	31-MAR-1998;	98US-0080107P.
PR	29-SEP-1998;	98US-0102240P.	PR	31-MAR-1998;	98US-0080194P.
PR	29-SEP-1998;	98US-0102330P.	PR	01-APR-1998;	98US-0080327P.
PR	29-SEP-1998;	98US-0102331P.	PR	01-APR-1998;	98US-0080333P.
PR	30-SEP-1998;	98US-0102487P.	PR	08-APR-1998;	98US-0081049P.
PR	30-SEP-1998;	98US-0102570P.	PR	08-APR-1998;	98US-0081070P.
PR	30-SEP-1998;	98US-0102571P.	PR	09-APR-1998;	98US-0081195P.
PR	01-OCT-1998;	98US-0102684P.	PR	09-APR-1998;	98US-0081838P.
PR	01-OCT-1998;	98US-0102687P.	PR	15-APR-1998;	98US-0081838P.
Query Match 60.5%; Score 49; DB 6; Length 199;					
Best Local Similarity 75.0%; Pred. No. 6.2;					
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	3 LVGPMSLWLC 14				
Db	184 LTGHSLMSWLC 195				
RESULT 29					
ABR65820					
ID	ABR65820 standard; protein; 199 AA.				
XX	AC ABR65820;				
XX	DT 05-AUG-2003 (first entry)				
XX	DE Human secreted polypeptide PRO19669, SEQ ID NO:588.				
XX	KW Human; PRO; secreted protein; transmembrane protein; extracellular domain; tumour necrosis factor-alpha; TNF-alpha;				
KW	chondrocyte; proliferation; differentiation; cartilage disorder;				
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;				
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;				
KW	liver; drug screening; transgenic animal; genetic analysis;				
KW	antiarthritic; vulnerrary; gene therapy.				
XX	OS Homo sapiens.				
XX	XX US2003036159-A1.				
XX	XX 20-FEB-2003.				
XX	XX 02-JUL-2002; 2002US-00188773.				
XX	XX 18-SEP-1997;				
PR	18-SEP-1997;	97US-0059263P.	PR	18-SEP-1997;	97US-0059266P.
PR	17-OCT-1997;	97US-0062250P.	PR	21-OCT-1997;	97US-0063486P.
PR	24-OCT-1997;	97US-0063120P.	PR	24-OCT-1997;	97US-0063120P.
PR	28-OCT-1997;	97US-0063540P.	PR	28-OCT-1997;	97US-0063541P.
PR	28-OCT-1997;	97US-0063541P.	PR	28-OCT-1997;	97US-0063544P.
PR	28-OCT-1997;	97US-0063564P.	PR	28-OCT-1997;	97US-0063564P.
PR	29-OCT-1997;	97US-0063734P.	PR	31-OCT-1997;	97US-0063870P.
PR	31-OCT-1997;	97US-0064103P.	PR	13-NOV-1997;	97US-0065311P.
PR	21-NOV-1997;	97US-00656120P.	PR	24-NOV-1997;	97US-0066466P.
PR	24-NOV-1997;	97US-0066772P.	PR	11-DEC-1997;	97US-0069335P.

```

PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090535P.
PR 25-JUN-1998; 98US-0090640P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 26-JUN-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090535P.
PR 25-JUN-1998; 98US-0090640P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 26-JUN-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPELSMWLTC 14
Db 184 LTGHSLMSWLVC 195

RESULT 30
ABU99760
ID ABU99760 standard; protein; 199 AA.
AC ABU99760;
XX 09-AUG-2003 (first entry)
DT Human secreted/transmembrane protein (PRO) #294.
DE Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX Homo sapiens.
XX US2003040070-A1.
XX 27-FEB-2003.
XX 27-JUN-2002; 2002US-00184627.
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063734P.
PR 29-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 31-OCT-1997; 97US-0065311P.
PR 13-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 20-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 27-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
```

PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083455P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084619P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0088027P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101688P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 25-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.

PR	01-OCT-1998;	98US-0102684P.	PR	21-APR-1998;	98US-0082568P.
PR	01-OCT-1998;	98US-0102687P.	PR	21-APR-1998;	98US-0082569P.
PR	02-OCT-1998;	98US-0102965P.	PR	22-APR-1998;	98US-0082704P.
PR	06-OCT-1998;	98US-0103258P.	PR	22-APR-1998;	98US-0082797P.
PR	06-OCT-1998;	98US-0103449P.	PR	28-APR-1998;	98US-0083322P.
PR	07-OCT-1998;	98US-00168978.	PR	29-APR-1998;	98US-0083455P.
Query Match 60.5%; Score 49; DB 6; Length 199;					
Best Local Similarity 75.0%; Pred. No. 6.2;					
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	3 LVGPSLSWLTC 14				
Db	184 LTGHSLSWLVC 195				
RESULT 31					
ABU82999					
ID	ABU82999 standard; protein; 199 AA.				
XX	AC ABU82999;				
DT	27-JUN-2003 (first entry)				
XX	Human PRO polypeptide #294.				
XX	Human; PRO polypeptide; secreted and transmembrane protein; tumour;				
KW	chromosome mapping; gene mapping; cytostatic.				
KW	Homo sapiens.				
OS	XX				
XX	US2003032113-A1.				
PN	13-FEB-2003.				
XX	20-JUN-2002; 2002US-00176911.				
PR	18-SEP-1997;	97US-0059263P.	PR	10-JUN-1998;	98US-0088738P.
PR	18-SEP-1997;	97US-0059266P.	PR	10-JUN-1998;	98US-0088740P.
PR	17-OCT-1997;	97US-0062250P.	PR	10-JUN-1998;	98US-0088811P.
PR	21-OCT-1997;	97US-0063486P.	PR	10-JUN-1998;	98US-0088824P.
PR	24-OCT-1997;	97US-0063120P.	PR	10-JUN-1998;	98US-0088825P.
PR	24-OCT-1997;	97US-0063121P.	PR	10-JUN-1998;	98US-0088826P.
PR	28-OCT-1997;	97US-0063540P.	PR	11-JUN-1998;	98US-0088861P.
PR	28-OCT-1997;	97US-0063541P.	PR	11-JUN-1998;	98US-0088863P.
PR	28-OCT-1997;	97US-0063544P.	PR	11-JUN-1998;	98US-0088867P.
PR	28-OCT-1997;	97US-0063564P.	PR	12-JUN-1998;	98US-0089090P.
PR	29-OCT-1997;	97US-0063734P.	PR	12-JUN-1998;	98US-0089105P.
PR	31-OCT-1997;	97US-0063870P.	PR	16-JUN-1998;	98US-0089512P.
PR	31-OCT-1997;	97US-0064103P.	PR	16-JUN-1998;	98US-0089514P.
PR	13-NOV-1997;	97US-0065311P.	PR	17-JUN-1998;	98US-0089538P.
PR	21-NOV-1997;	97US-0066120P.	PR	17-JUN-1998;	98US-0089598P.
PR	24-NOV-1997;	97US-0066466P.	PR	17-JUN-1998;	98US-0089653P.
PR	24-NOV-1997;	97US-0066772P.	PR	18-JUN-1998;	98US-0089908P.
PR	11-DEC-1997;	97US-0069333P.	PR	19-JUN-1998;	98US-0089952P.
PR	12-DEC-1997;	97US-0069425P.	PR	22-JUN-1998;	98US-0090252P.
PR	17-DEC-1997;	97US-0069870P.	PR	22-JUN-1998;	98US-0090254P.
PR	18-DEC-1997;	97US-0068017P.	PR	24-JUN-1998;	98US-0090429P.
PR	10-MAR-1998;	98US-0077450P.	PR	24-JUN-1998;	98US-0090435P.
PR	11-MAR-1998;	98US-0077632P.	PR	24-JUN-1998;	98US-0090444P.
PR	11-MAR-1998;	98US-0077649P.	PR	24-JUN-1998;	98US-0090461P.
PR	20-MAR-1998;	98US-0078886P.	PR	24-JUN-1998;	98US-0090535P.
PR	20-MAR-1998;	98US-0078939P.	PR	24-JUN-1998;	98US-0090538P.
PR	27-MAR-1998;	98US-0079664P.	PR	25-JUN-1998;	98US-0090540P.
PR	27-MAR-1998;	98US-0079786P.	PR	25-JUN-1998;	98US-0090676P.
PR	31-MAR-1998;	98US-0080107P.	PR	25-JUN-1998;	98US-0090688P.
PR	31-MAR-1998;	98US-0080194P.	PR	25-JUN-1998;	98US-0090689P.
PR	01-APR-1998;	98US-0080327P.	PR	25-JUN-1998;	98US-0090694P.
PR	01-APR-1998;	98US-0080333P.	PR	25-JUN-1998;	98US-0090695P.
PR	08-APR-1998;	98US-0081049P.	PR	25-JUN-1998;	98US-0090696P.
PR	08-APR-1998;	98US-0081070P.	PR	26-JUN-1998;	98US-00105413.
PR	09-APR-1998;	98US-0081195P.	PR	26-JUN-1998;	98US-0090862P.
PR	15-APR-1998;	98US-0081838P.			

```
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 02-JUL-1998; 98US-0091633P.
PR 04-AUG-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097282P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100622P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101777P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102027P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-0103395P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTC 14
Db 184 LTGHSLMSWLVC 195

RESULT 32
ABU90120
ID ABU90120 standard; protein; 199 AA.
XX AC ABU90120;
XX DT 11-AUG-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO19669.
XX KW Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;
XX KW chondrocyte differentiation; tumour necrosis factor-alpha release;
XX KW affinity purification.
XX OS Homo sapiens.
XX PN US2003036147-A1.
XX PD 20-FEB-2003.
XX PF 02-JUL-2002; 2002US-00187741.
XX PR 18-SEP-1997; 97US-0059263P.
XX PR 18-SEP-1997; 97US-0059486P.
XX PR 21-OCT-1997; 97US-0062250P.
XX PR 24-OCT-1997; 97US-0063486P.
XX PR 24-OCT-1997; 97US-0063120P.
XX PR 28-OCT-1997; 97US-0063121P.
XX PR 28-OCT-1997; 97US-0063540P.
XX PR 28-OCT-1997; 97US-0063541P.
XX PR 28-OCT-1997; 97US-0063544P.
XX PR 29-OCT-1997; 97US-0063564P.
XX PR 31-OCT-1997; 97US-0063734P.
XX PR 31-OCT-1997; 97US-0063870P.
XX PR 13-NOV-1997; 97US-0064103P.
XX PR 21-NOV-1997; 97US-0066120P.
XX PR 24-NOV-1997; 97US-0066466P.
XX PR 11-DEC-1997; 97US-0066772P.
XX PR 12-DEC-1997; 97US-0069335P.
XX PR 17-DEC-1997; 97US-0069425P.
XX PR 18-DEC-1997; 97US-0069870P.
XX PR 10-MAR-1998; 98US-0077450P.
XX PR 11-MAR-1998; 98US-0077632P.
XX PR 20-MAR-1998; 98US-0077649P.
XX PR 20-MAR-1998; 98US-0078886P.
XX PR 27-MAR-1998; 98US-0078939P.
XX PR 27-MAR-1998; 98US-0079664P.
XX PR 31-MAR-1998; 98US-0080107P.
XX PR 31-MAR-1998; 98US-0080194P.
XX PR 01-APR-1998; 98US-0080327P.
XX PR 08-APR-1998; 98US-0080333P.
XX PR 08-APR-1998; 98US-0081049P.
XX PR 09-APR-1998; 98US-0081070P.
XX PR 15-APR-1998; 98US-0081195P.
XX PR 21-APR-1998; 98US-0081838P.
XX PR 21-APR-1998; 98US-0082568P.
XX PR 22-APR-1998; 98US-0082569P.
XX PR 22-APR-1998; 98US-0082704P.
XX PR 28-APR-1998; 98US-0082797P.
XX PR 29-APR-1998; 98US-0083322P.
XX PR 29-APR-1998; 98US-0083495P.
```

PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087603P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00905413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.

PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLSMWLTC 14
Db 184 LTGHSLSMWLVC 195

RESULT 33
ABR68369
ID ABR68369 standard; protein; 199 AA.
XX
AC ABR68369;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003027264-A1.
XX
PD 06-FEB-2003.
XX
PF 18-JUN-2002; 2002US-00174579.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0082250P.
PR 21-OCT-1997; 97US-0083486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0083544P.
PR 28-OCT-1997; 97US-0083564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069435P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 29-APR-1998; 98US-0083560P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 15-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0088909P.
PR 12-JUN-1998; 98US-00889105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.

PR	24-JUL-1998;	98US-0094006P.	XX	25-JUL-2003	(first entry)
PR	04-AUG-1998;	98US-0095282P.	DT		
PR	10-AUG-1998;	98US-0095998P.	XX		Novel human secreted and transmembrane protein PRO19669.
PR	10-AUG-1998;	98US-0096012P.	DE		
PR	17-AUG-1998;	98US-0096757P.	XX		Human; secreted and transmembrane protein; PRO; transgenic animal;
PR	17-AUG-1998;	98US-0096766P.	KW		knockout; chromosome identification; tissue typing; tumour;
PR	17-AUG-1998;	98US-0096867P.	KW		chondrocyte proliferation; chondrocyte differentiation;
PR	17-AUG-1998;	98US-0096881P.	KW		tumor necrosis factor-alpha release stimulator.
PR	17-AUG-1998;	98US-0096897P.	XX		
PR	18-AUG-1998;	98US-0096949P.	OS		Homo sapiens.
PR	18-AUG-1998;	98US-0096959P.	XX		
PR	18-AUG-1998;	98US-0097022P.	XX		US2003036144-A1.
PR	26-AUG-1998;	98US-0097952P.	XX		
PR	26-AUG-1998;	98US-0097954P.	PD		20-FEB-2003.
PR	26-AUG-1998;	98US-0097955P.	XX		
PR	26-AUG-1998;	98US-0097971P.	XX		01-JUL-2002; 2002US-00187601.
PR	26-AUG-1998;	98US-0097974P.	XX		
PR	01-SEP-1998;	98US-0098014P.	XX		
PR	01-SEP-1998;	98US-0098016P.	PR	18-SEP-1997;	97US-0059263P.
PR	01-SEP-1998;	98US-0098073P.	PR	18-SEP-1997;	97US-0059266P.
PR	02-SEP-1998;	98US-0098803P.	PR	17-OCT-1997;	97US-0062250P.
PR	02-SEP-1998;	98US-0098821P.	PR	21-OCT-1997;	97US-0063486P.
PR	02-SEP-1998;	98US-0098843P.	PR	24-OCT-1997;	97US-0063120P.
PR	09-SEP-1998;	98US-0099602P.	PR	24-OCT-1997;	97US-0063121P.
PR	10-SEP-1998;	98US-0099741P.	PR	28-OCT-1997;	97US-0063540P.
PR	10-SEP-1998;	98US-0099754P.	PR	28-OCT-1997;	97US-0063541P.
PR	10-SEP-1998;	98US-0099763P.	PR	28-OCT-1997;	97US-0063544P.
PR	15-SEP-1998;	98US-0100388P.	PR	28-OCT-1997;	97US-0063564P.
PR	16-SEP-1998;	98US-0100662P.	PR	28-OCT-1997;	97US-0063734P.
PR	16-SEP-1998;	98US-0100664P.	PR	31-OCT-1997;	97US-0063870P.
PR	16-SEP-1998;	98US-0101751P.	PR	31-OCT-1997;	97US-0064103P.
PR	16-SEP-1998;	98US-0101933P.	PR	13-NOV-1997;	97US-0065311P.
PR	17-SEP-1998;	98US-0100683P.	PR	21-NOV-1997;	97US-0066120P.
PR	17-SEP-1998;	98US-0100684P.	PR	24-NOV-1997;	97US-0066466P.
PR	17-SEP-1998;	98US-0100919P.	PR	24-NOV-1997;	97US-0066772P.
PR	17-SEP-1998;	98US-0100930P.	PR	11-DEC-1997;	97US-0069335P.
PR	18-SEP-1998;	98US-0100849P.	PR	12-DEC-1997;	97US-0069425P.
PR	18-SEP-1998;	98US-0101014P.	PR	17-DEC-1997;	97US-0069870P.
PR	18-SEP-1998;	98US-0101068P.	PR	18-DEC-1997;	97US-0068017P.
PR	23-SEP-1998;	98US-0101471P.	PR	10-MAR-1998;	98US-0077450P.
PR	23-SEP-1998;	98US-0101472P.	PR	11-MAR-1998;	98US-0077632P.
PR	23-SEP-1998;	98US-0101473P.	PR	11-MAR-1998;	98US-0077649P.
PR	23-SEP-1998;	98US-0101477P.	PR	20-MAR-1998;	98US-0078886P.
PR	24-SEP-1998;	98US-0101738P.	PR	20-MAR-1998;	98US-0078939P.
PR	24-SEP-1998;	98US-0101739P.	PR	27-MAR-1998;	98US-0079664P.
PR	24-SEP-1998;	98US-0101743P.	PR	27-MAR-1998;	98US-0079786P.
PR	24-SEP-1998;	98US-0101922P.	PR	31-MAR-1998;	98US-0080107P.
PR	25-SEP-1998;	98US-0101786P.	PR	31-MAR-1998;	98US-0080194P.
PR	25-SEP-1998;	98US-0102207P.	PR	01-APR-1998;	98US-0080327P.
PR	29-SEP-1998;	98US-0102207P.	PR	01-APR-1998;	98US-0080333P.
PR	29-SEP-1998;	98US-0102240P.	PR	08-APR-1998;	98US-0081049P.
PR	29-SEP-1998;	98US-0102330P.	PR	08-APR-1998;	98US-0081070P.
PR	29-SEP-1998;	98US-0102331P.	PR	09-APR-1998;	98US-0081195P.
PR	29-SEP-1998;	98US-0102331P.	PR	15-APR-1998;	98US-0081838P.
PR	30-SEP-1998;	98US-0102487P.	PR	21-APR-1998;	98US-0082568P.
PR	30-SEP-1998;	98US-0102570P.	PR	21-APR-1998;	98US-0082569P.
PR	30-SEP-1998;	98US-0102571P.	PR	22-APR-1998;	98US-0082704P.
PR	01-OCT-1998;	98US-0102684P.	PR	22-APR-1998;	98US-0082797P.
PR	01-OCT-1998;	98US-0102687P.	PR	28-APR-1998;	98US-0083322P.
			PR	29-APR-1998;	98US-0083495P.
			PR	29-APR-1998;	98US-0083496P.
			PR	29-APR-1998;	98US-0083499P.
			PR	29-APR-1998;	98US-0083559P.
			PR	05-MAY-1998;	98US-0084366P.
			PR	06-MAY-1998;	98US-0084414P.
			PR	07-MAY-1998;	98US-0084639P.
			PR	07-MAY-1998;	98US-0084640P.
			PR	07-MAY-1998;	98US-0084643P.
			PR	15-MAY-1998;	98US-0085579P.
			PR	15-MAY-1998;	98US-0085580P.
			PR	15-MAY-1998;	98US-0085582P.
			PR	15-MAY-1998;	98US-0085700P.
			PR	18-MAY-1998;	98US-0086023P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLMSWLC 14
DB 184 LTGHSLSWLC 195

RESULT 34
ABU96422
ID ABU96422 standard; protein; 199 AA.
XX
AC ABU96422;

PR 22-MAY-1998; 98US-00863392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-008787P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088871P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089152P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090541P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.

PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 03-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 30-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGFSLSWLTLC 14
| | | | | | | |
Db 184 LTGHSLMSWLVC 195

RESULT 35

ABU92853
ID ABU92853 standard; protein; 199 AA.

XX
AC ABU92853;

XX
DT 18-JUL-2003 (first entry)

XX
DE Human secreted/transmembrane protein (PRO) #294.

XX
KW Human; secreted protein; transmembrane protein; PRO; tumour;
proliferation; differentiation; chondrocyte cell; TNF-alpha;

KW	tumour necrosis factor-alpha; gene therapy.	PR	04-JUN-1998;	98US-0088028P.
XX		PR	04-JUN-1998;	98US-0088029P.
OS	Homo sapiens.	PR	04-JUN-1998;	98US-0088033P.
XX		PR	04-JUN-1998;	98US-0088326P.
PN	US2003036149-A1.	PR	05-JUN-1998;	98US-0088167P.
XX		PR	05-JUN-1998;	98US-0088202P.
XX		PR	05-JUN-1998;	98US-0088212P.
PD	20-FEB-2003.	PR	05-JUN-1998;	98US-0088217P.
XX		PR	09-JUN-1998;	98US-0088655P.
PF	02-JUL-2002; 2002US-00187746.	PR	10-JUN-1998;	98US-0088722P.
XX		PR	10-JUN-1998;	98US-0088738P.
XX	18-SEP-1997;	PR	10-JUN-1998;	98US-0088740P.
PR	97US-0059263P.	PR	10-JUN-1998;	98US-0088811P.
PR	18-SEP-1997;	PR	10-JUN-1998;	98US-0088824P.
PR	97US-0059266P.	PR	10-JUN-1998;	98US-0088825P.
PR	17-OCT-1997;	PR	10-JUN-1998;	98US-0088826P.
PR	97US-0063486P.	PR	10-JUN-1998;	98US-0088861P.
PR	21-OCT-1997;	PR	10-JUN-1998;	98US-0088863P.
PR	97US-0063120P.	PR	10-JUN-1998;	98US-0088865P.
PR	24-OCT-1997;	PR	10-JUN-1998;	98US-0088866P.
PR	97US-0063121P.	PR	11-JUN-1998;	98US-0088867P.
PR	28-OCT-1997;	PR	11-JUN-1998;	98US-0088876P.
PR	97US-0063540P.	PR	11-JUN-1998;	98US-0089090P.
PR	28-OCT-1997;	PR	12-JUN-1998;	98US-0089105P.
PR	97US-0063544P.	PR	12-JUN-1998;	98US-0089512P.
PR	28-OCT-1997;	PR	16-JUN-1998;	98US-0089514P.
PR	97US-0063564P.	PR	16-JUN-1998;	98US-0089538P.
PR	29-OCT-1997;	PR	17-JUN-1998;	98US-0089539P.
PR	97US-0063734P.	PR	17-JUN-1998;	98US-0089653P.
PR	31-OCT-1997;	PR	17-JUN-1998;	98US-0089908P.
PR	97US-0064103P.	PR	18-JUN-1998;	98US-0089952P.
PR	13-NOV-1997;	PR	18-JUN-1998;	98US-0090246P.
PR	97US-0065311P.	PR	19-JUN-1998;	98US-0090252P.
PR	21-NOV-1997;	PR	22-JUN-1998;	98US-0090254P.
PR	97US-0066468P.	PR	22-JUN-1998;	98US-0090429P.
PR	24-NOV-1997;	PR	24-JUN-1998;	98US-0090435P.
PR	97US-0066772P.	PR	24-JUN-1998;	98US-0090444P.
PR	11-DEC-1997;	PR	24-JUN-1998;	98US-0090461P.
PR	97US-0069333P.	PR	24-JUN-1998;	98US-0090535P.
PR	12-DEC-1997;	PR	24-JUN-1998;	98US-0090540P.
PR	97US-0069425P.	PR	25-JUN-1998;	98US-0090676P.
PR	17-DEC-1997;	PR	25-JUN-1998;	98US-0090678P.
PR	97US-0069870P.	PR	25-JUN-1998;	98US-0090688P.
PR	18-DEC-1997;	PR	25-JUN-1998;	98US-0090690P.
PR	97US-0068017P.	PR	25-JUN-1998;	98US-0090695P.
PR	10-MAR-1998;	PR	25-JUN-1998;	98US-0090696P.
PR	98US-0077450P.	PR	25-JUN-1998;	98US-0090699P.
PR	11-MAR-1998;	PR	26-JUN-1998;	98US-0090862P.
PR	98US-0077632P.	PR	26-JUN-1998;	98US-0090863P.
PR	11-MAR-1998;	PR	26-JUN-1998;	98US-0090863P.
PR	98US-0077649P.	PR	26-JUN-1998;	98US-0091010P.
PR	20-MAR-1998;	PR	26-JUN-1998;	98US-0091359P.
PR	98US-0078886P.	PR	26-JUN-1998;	98US-0091544P.
PR	20-MAR-1998;	PR	26-JUN-1998;	98US-0091478P.
PR	98US-0079664P.	PR	26-JUN-1998;	98US-0091486P.
PR	27-MAR-1998;	PR	26-JUN-1998;	98US-0091626P.
PR	98US-0079786P.	PR	26-JUN-1998;	98US-0091628P.
PR	27-MAR-1998;	PR	26-JUN-1998;	98US-0091632P.
PR	98US-0080107P.	PR	26-JUN-1998;	98US-0091632P.
PR	31-MAR-1998;	PR	26-JUN-1998;	98US-0094006P.
PR	98US-0080194P.	PR	26-JUN-1998;	98US-0095282P.
PR	01-APR-1998;	PR	26-JUN-1998;	98US-0095598P.
PR	98US-0080327P.	PR	26-JUN-1998;	98US-0096012P.
PR	01-APR-1998;	PR	26-JUN-1998;	98US-0096757P.
PR	98US-0080333P.	PR	26-JUN-1998;	98US-0096766P.
PR	08-APR-1998;	PR	26-JUN-1998;	98US-0096867P.
PR	98US-0081049P.	PR	26-JUN-1998;	98US-0096891P.
PR	08-APR-1998;	PR	26-JUN-1998;	98US-0096897P.
PR	98US-0081070P.	PR	26-JUN-1998;	98US-0096949P.
PR	09-APR-1998;	PR	26-JUN-1998;	98US-0097022P.
PR	98US-0081195P.	PR	26-JUN-1998;	98US-0097952P.
PR	15-APR-1998;	PR	26-JUN-1998;	98US-0097954P.
PR	98US-0081838P.	PR	26-JUN-1998;	98US-0097955P.
PR	21-APR-1998;	PR	26-JUN-1998;	98US-0097971P.
PR	98US-0082568P.	PR	26-JUN-1998;	98US-0097974P.
PR	08US-0082569P.	PR	26-JUN-1998;	98US-0098014P.
PR	21-APR-1998;	PR	26-JUN-1998;	
PR	98US-0082704P.	PR	26-JUN-1998;	
PR	22-APR-1998;	PR	26-JUN-1998;	
PR	98US-0082797P.	PR	26-JUN-1998;	
PR	22-APR-1998;	PR	26-JUN-1998;	
PR	98US-0083322P.	PR	26-JUN-1998;	
PR	28-APR-1998;	PR	26-JUN-1998;	
PR	98US-0083495P.	PR	26-JUN-1998;	
PR	29-APR-1998;	PR	26-JUN-1998;	
PR	98US-0083496P.	PR	26-JUN-1998;	
PR	29-APR-1998;	PR	26-JUN-1998;	
PR	98US-0083499P.	PR	26-JUN-1998;	
PR	29-APR-1998;	PR	26-JUN-1998;	
PR	98US-0083559P.	PR	26-JUN-1998;	
PR	05-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0084366P.	PR	26-JUN-1998;	
PR	06-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0084414P.	PR	26-JUN-1998;	
PR	07-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0084639P.	PR	26-JUN-1998;	
PR	07-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0084640P.	PR	26-JUN-1998;	
PR	07-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0084643P.	PR	26-JUN-1998;	
PR	15-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0085580P.	PR	26-JUN-1998;	
PR	15-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0085582P.	PR	26-JUN-1998;	
PR	15-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0085700P.	PR	26-JUN-1998;	
PR	18-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0086023P.	PR	26-JUN-1998;	
PR	22-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0086392P.	PR	26-JUN-1998;	
PR	28-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0086486P.	PR	26-JUN-1998;	
PR	28-MAY-1998;	PR	26-JUN-1998;	
PR	98US-0087098P.	PR	26-JUN-1998;	
PR	02-JUN-1998;	PR	26-JUN-1998;	
PR	98US-0087609P.	PR	26-JUN-1998;	
PR	02-JUN-1998;	PR	26-JUN-1998;	
PR	98US-0087759P.	PR	26-JUN-1998;	
PR	03-JUN-1998;	PR	26-JUN-1998;	
PR	98US-0087827P.	PR	26-JUN-1998;	
PR	04-JUN-1998;	PR	26-JUN-1998;	
PR	98US-0088025P.	PR	26-JUN-1998;	

```
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 03-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98MO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102470P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSTMSLTC 14
Db 184 LTGHSLSLVC 195

RESULT 36
ABO08930
ID ABO08930 standard; protein; 199 AA.
XX
AC ABO08930;
XX
DT 17-AUG-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #294.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
FN US2003044923-A1.
XX
```

```
PD 06-MAR-2003.
XX
PF 24-JUN-2002; 2002US-00179522.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 19-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 05-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
```

PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089582P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090433P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-00999602P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089582P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090433P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-00999602P.

PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 10-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102310P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPGLMSWLTC 14
| | | | | | | | | |
Db 184 LTGHSLSMLVC 195
RESULT 37
ABO02982
ID ABO02982 standard; protein; 199 AA.
XX ABO02982;
XX 09-AUG-2003 (first entry)
XX Human secreted/transmembrane protein (PRO) #294.
XX Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing; adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX Homo sapiens.
XX US2003040062-A1.
XX 27-FEB-2003.
XX 25-JUN-2002; 2002US-00180545.
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.

PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0065933P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083455P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085583P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.

PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.


```
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100652P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98MO-U8019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100652P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98MO-U8019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLSMLTLC 14
Db 184 LTGHSLSMLVC 195
RESULT 39
ABR94898
ID ABR94898 standard; protein; 199 AA.
XX
AC ABR94898;
XX
DT 13-SEP-2003 (first entry)
XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
liver; drug screening; transgenic animal; genetic analysis;
antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003044926-A1.
XX
PD 06-MAR-2003.
XX
PF 26-JUN-2002; 2002US-00183015.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
```


PR	17-JUN-1998;	98US-0089538P;
PR	17-JUN-1998;	98US-0089656P;
PR	18-JUN-1998;	98US-0089608P;
PR	19-JUN-1998;	98US-0089952P;
PR	22-JUN-1998;	98US-0090242P;
PR	22-JUN-1998;	98US-0090254P;
PR	22-JUN-1998;	98US-0090252P;
PR	24-JUN-1998;	98US-0090429P;
PR	24-JUN-1998;	98US-0090435P;
PR	24-JUN-1998;	98US-0090444P;
PR	24-JUN-1998;	98US-0090461P;
PR	24-JUN-1998;	98US-0090535P;
PR	24-JUN-1998;	98US-0090540P;
PR	25-JUN-1998;	98US-0090676P;
PR	25-JUN-1998;	98US-0090678P;
PR	25-JUN-1998;	98US-0090680P;
PR	25-JUN-1998;	98US-0090686P;
PR	25-JUN-1998;	98US-0090694P;
PR	25-JUN-1998;	98US-0090695P;
PR	25-JUN-1998;	98US-0090696P;
PR	25-JUN-1998;	98US-0090697P;
PR	26-JUN-1998;	98US-0090861P;
PR	26-JUN-1998;	98US-0090862P;
PR	26-JUN-1998;	98US-0090863P;
PR	26-JUN-1998;	98US-0091101P;
PR	01-JUL-1998;	98US-0091135P;
PR	01-JUL-1998;	98US-0091544P;
PR	02-JUL-1998;	98US-0091478P;
PR	02-JUL-1998;	98US-0091486P;
PR	02-JUL-1998;	98US-0091626P;
PR	02-JUL-1998;	98US-0091628P;
PR	02-JUL-1998;	98US-0091632P;
PR	24-JUL-1998;	98US-0094006P;
PR	24-JUL-1998;	98US-0095282P;
PR	10-AUG-1998;	98US-0095998P;
PR	10-AUG-1998;	98US-0096012P;
PR	17-AUG-1998;	98US-0096757P;
PR	17-AUG-1998;	98US-0096766P;
PR	17-AUG-1998;	98US-0096867P;
PR	17-AUG-1998;	98US-0096891P;
PR	17-AUG-1998;	98US-0096897P;
PR	18-AUG-1998;	98US-0096949P;
PR	18-AUG-1998;	98US-0096959P;
PR	18-AUG-1998;	98US-0097022P;
PR	26-AUG-1998;	98US-0097952P;
PR	26-AUG-1998;	98US-0097954P;
PR	26-AUG-1998;	98US-0097955P;
PR	26-AUG-1998;	98US-0097971P;
PR	26-AUG-1998;	98US-0097974P;
PR	26-AUG-1998;	98US-0098014P;
PR	01-SEP-1998;	98US-0098716P;
PR	01-SEP-1998;	98US-0098723P;
PR	02-SEP-1998;	98US-0098803P;
PR	02-SEP-1998;	98US-0098821P;
PR	02-SEP-1998;	98US-0098812P;
PR	15-SEP-1998;	98US-0100388P;
PR	16-SEP-1998;	98US-0100562P;
PR	16-SEP-1998;	98US-0100564P;
PR	16-SEP-1998;	98US-0101751P;
PR	16-SEP-1998;	98US-0101751P;
PR	17-SEP-1998;	98US-0100693P;
PR	17-SEP-1998;	98US-0100684P;
PR	17-SEP-1998;	98US-0100919P;
PR	17-SEP-1998;	98US-0100930P;
PR	18-SEP-1998;	98US-0100849P;
PR	18-SEP-1998;	98US-0101014P;
PR	23-SEP-1998;	98US-0101068P;
PR	23-SEP-1998;	98US-0101471P;
PR	23-SEP-1998;	98US-0101472P;

PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.
 PR 25-SEP-1998; 98US-0101786P.
 PR 25-SEP-1998; 98US-0102207P.
 PR 25-SEP-1998; 98US-0102240P.
 PR 25-SEP-1998; 98US-0102330P.
 PR 25-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;

Best Local Similarity 75.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGSLMSWLTC 14

Db 184 LTGHSLSWLVC 195

RESULT 40

ABU85871

ID ABU85871 standard; protein; 199 AA.

XX AC ABU85871;

XX DT 02-JUL-2003 (first entry)

XX DE Human PRO polypeptide #294.

XX KW Human; PRO; secreted polypeptide; transmembrane polypeptide;
 tumour necrosis factor alpha; TNF-alpha; chondrocyte cell; tumour;
 cytostatic.

XX OS Homo sapiens.

XX PN US2003036140-A1.

XX PD 20-FEB-2003.

XX PF 01-JUL-2002; 2002US-00187588.

XX PR 26-JUN-1998; 98US-00105413.

XX PR 16-SEP-1998; 98WO-US019330.

XX PR 07-OCT-1998; 98US-00168978.

XX PR 06-NOV-1998; 98WO-US021141.

XX PR 01-DEC-1998; 98US-00187368.

XX PR 07-DEC-1998; 98WO-US025108.

XX PR 03-MAR-1999; 98US-00254311.

XX PR 08-MAR-1999; 99WO-US005028.

XX PR 14-MAY-1999; 99US-00311832.

XX PR 02-JUN-1999; 99WO-US010733.

XX PR 25-AUG-1999; 99WO-US012252.

XX PR 25-AUG-1999; 99US-00380137.

XX PR 25-AUG-1999; 99US-00380138.

XX PR 25-AUG-1999; 99US-00380139.

XX PR 01-SEP-1999; 99WO-US020111.

XX PR 15-SEP-1999; 99WO-US021090.

XX PR 18-OCT-1999; 99US-00403297.

XX PR 12-NOV-1999; 99US-00423844.

XX PR 01-DEC-1999; 99WO-US028301.

XX PR 30-DEC-1999; 99WO-US028551.

XX PR 05-JAN-2000; 99WO-US031374.

XX PR 18-FEB-2000; 2000WO-US000219.

XX PR 18-FEB-2000; 2000WO-US004341.

PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 22-AUG-2000; 2000US-00644848.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 18-SEP-2000; 2000US-00664610.
 PR 18-SEP-2000; 2000US-00665350.
 PR 08-NOV-2000; 2000US-00709238.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 28-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 22-MAR-2001; 2001US-00816744.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 25-MAY-2001; 2001US-00866028.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 30-JUL-2001; 2001US-00918585.
 PR 06-AUG-2001; 2001US-00924419.
 PR 13-AUG-2001; 2001US-00929404.
 PR 16-AUG-2001; 2001US-00931836.
 PR 28-AUG-2001; 2001US-00941992.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 04-SEP-2001; 2001US-00946374.
 PR 15-JAN-2002; 2002US-00052586.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;

Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-332028/31.

N-PSDB; ACA73064.

Three hundred and five nucleic acids encoding PRO polypeptides, useful for the manufacture of a medicament for diagnosing or treating tumor.

Claim 11; Fig 588; 707pp; English.

The invention relates to human PRO polypeptides (secreted and transmembrane polypeptides) and the PRO polynucleotides encoding them. The invention also relates to a method for stimulating the release of tumour necrosis factor alpha (TNF-alpha) from human blood by contacting the blood with a sequence of the invention, a method for stimulating the proliferation or differentiation of chondrocyte cells by contacting the cells with a PRO polypeptide and a method for detecting the presence of a tumour in a mammal. The polypeptides and polynucleotides are useful for the manufacture of a medicament for diagnosing or treating a tumour in a mammal. Sequences ABU85578-ABU85882 represent human PRO polypeptides of the invention. Note: The sequence data for this patent is also available in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 199 AA;

Query Match 60.5%; Score 49; DB 6; Length 199;

Best Local Similarity 75.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTLC 14
Db 184 LTGHSLMSWLTVC 195

RESULT 41
ABU99031
ID ABU99031 standard; protein; 199 AA.
XX
AC ABU99031;
XX
DT 01-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO19669.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy;
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release;
KW tumour necrosis factor alpha release; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;
KW bioreactor.
XX
OS Homo sapiens.
XX
FN US2003013153-A1.
XX PD
XX PD 16-JAN-2003.
XX PF 19-JUN-2002; 2002US-00175737.
XX PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 31-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069333P.
PR 12-DEC-1997; 97US-0069423P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 05-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.

```
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98MO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 25-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 LVGPSLMSWLTG 14
Db 184 LTGHSLSWLVG 195

RESULT 42
ABU98246
```

ABU98246 standard; protein; 199 AA.

ABU98246;

30-JUL-2003 (first entry)

Novel human secreted and transmembrane protein PRO19669.

Human; secreted and transmembrane protein; PRO; cytostatic; gene therapy; chondrocyte stimulator; tumour; adrenal tumour; lung tumour; colon tumour; breast tumour; prostate tumour; rectal tumour; cervical tumour; liver tumour; chromosome identification.

Homo sapiens.

US2003017544-A1.

23-JAN-2003.

21-JUN-2002; 2002US-00176915.

18-SEP-1997; 97US-0059263P.
18-SEP-1997; 97US-0059266P.
17-OCT-1997; 97US-0062250P.
21-OCT-1997; 97US-0063486P.
24-OCT-1997; 97US-0063120P.
24-OCT-1997; 97US-0063121P.
28-OCT-1997; 97US-0063540P.
28-OCT-1997; 97US-0063541P.
28-OCT-1997; 97US-0063544P.
28-OCT-1997; 97US-0063564P.
28-OCT-1997; 97US-0063734P.
31-OCT-1997; 97US-0063870P.
31-OCT-1997; 97US-0064103P.
13-NOV-1997; 97US-0065311P.
21-NOV-1997; 97US-0066120P.
24-NOV-1997; 97US-0066466P.
24-NOV-1997; 97US-0066772P.
11-DEC-1997; 97US-0069335P.
17-DEC-1997; 97US-0069425P.
18-DEC-1997; 97US-0069870P.
18-DEC-1997; 97US-0068017P.
10-MAR-1998; 98US-0077450P.
11-MAR-1998; 98US-0077632P.
11-MAR-1998; 98US-0077649P.
20-MAR-1998; 98US-0078886P.
27-MAR-1998; 98US-0079664P.
27-MAR-1998; 98US-0079786P.
31-MAR-1998; 98US-0080107P.
31-MAR-1998; 98US-0080194P.
01-APR-1998; 98US-0080327P.
08-APR-1998; 98US-0080333P.
08-APR-1998; 98US-0081049P.
08-APR-1998; 98US-0081070P.
09-APR-1998; 98US-0081195P.
15-APR-1998; 98US-0081838P.
21-APR-1998; 98US-0082568P.
21-APR-1998; 98US-0082569P.
22-APR-1998; 98US-0082704P.
22-APR-1998; 98US-0082797P.
28-APR-1998; 98US-0083322P.
29-APR-1998; 98US-0083495P.
29-APR-1998; 98US-0083496P.
29-APR-1998; 98US-0083499P.
05-MAY-1998; 98US-0083559P.
06-MAY-1998; 98US-0084366P.
06-MAY-1998; 98US-0084414P.
07-MAY-1998; 98US-0084639P.
07-MAY-1998; 98US-0084640P.
15-MAY-1998; 98US-0084643P.
15-MAY-1998; 98US-0085579P.
15-MAY-1998; 98US-0085580P.

PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088036P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.

PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 12-SEP-1998; 98US-0100388P.
PR 12-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-010207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGFSLSWLTC 14
| | | | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 43
ABU91952
ID ABU91952 standard; protein; 199 AA.
XX
AC ABU91952;
XX
DT 11-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO19669.

XX Human; gene therapy; chromosome identification; tissue typing.
KW Homo sapiens.
XX
OS
PN US2003027277-A1.
XX
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00178985.
XX
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 18-SEP-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 21-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 03-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.

PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090896P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095598P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.

XX	OS	Homo sapiens.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
----	----	---------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088226P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 05-JUN-1998; 98US-0088255P.
PR 09-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089553P.
PR 18-JUN-1998; 98US-0089808P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090576P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0090866P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 01-SEP-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101933P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLT 14
| | | | | | | |
Db 184 LTGHSLSWLVC 195

RESULT 45

ABU86486
ID ABU86486 standard; protein; 199 AA.

XX AC ABU86486;

XX DT 01-JUL-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) #294.

XX DE Human; immunogen; secreted protein; transmembrane protein; PRO; tumour;
XX KW proliferation; differentiation; chondrocyte cells;
XX KW tumour necrosis factor-alpha; TNF-alpha; blood; gene therapy.

XX OS Homo sapiens.

XX PN US2003036146-A1.

XX PD 20-FEB-2003.

XX PF 02-JUL-2002; 2002US-00187603.

XX PR 26-JUN-1998; 98US-00105413.

16-SEP-1998; 98WO-US019330.
07-OCT-1998; 98US-00168978.
07-OCT-1998; 98WO-US021141.
06-NOV-1998; 98US-00187368.
01-DEC-1998; 98WO-US025108.
07-DEC-1998; 98US-00202054.
03-MAR-1999; 99US-00254311.
08-MAR-1999; 99WO-US005028.
14-MAY-1999; 99US-00311832.
12-JUN-1999; 99WO-US010733.
25-AUG-1999; 99US-00380137.
25-AUG-1999; 99US-00380138.
25-AUG-1999; 99US-00380139.
25-AUG-1999; 99US-00380142.
01-SEP-1999; 99WO-US020111.
15-SEP-1999; 99WO-US021090.
18-OCT-1999; 99US-00403297.
12-NOV-1999; 99US-00423844.
01-DEC-1999; 99WO-US028301.
02-DEC-1999; 99WO-US028551.
30-DEC-1999; 99WO-US031274.
05-JAN-2000; 2000WO-US000219.
18-FEB-2000; 2000WO-US004341.
18-FEB-2000; 2000WO-US004342.
22-FEB-2000; 2000WO-US004414.
24-FEB-2000; 2000WO-US005004.
01-MAR-2000; 2000WO-US005601.
02-MAR-2000; 2000WO-US005841.
15-MAR-2000; 2000WO-US006884.
30-MAR-2000; 2000WO-US008439.
17-MAY-2000; 2000WO-US013705.
22-MAY-2000; 2000WO-US014042.
30-MAY-2000; 2000WO-US014941.
02-JUN-2000; 2000WO-US015264.
28-JUL-2000; 2000WO-US020710.
24-AUG-2000; 2000US-00844848.
24-AUG-2000; 2000WO-US023328.
18-SEP-2000; 2000US-00664610.
18-SEP-2000; 2000US-00665350.
08-NOV-2000; 2000US-00709238.
08-NOV-2000; 2000WO-US030952.
01-DEC-2000; 2000WO-US032678.
20-DEC-2000; 2000US-00747259.
28-FEB-2001; 2000WO-US034956.
22-MAR-2001; 2001WO-US006520.
10-MAY-2001; 2001US-00816744.
10-MAY-2001; 2001US-00854208.
25-MAY-2001; 2001US-00854280.
01-JUN-2001; 2001WO-US017800.
05-JUN-2001; 2001US-00874503.
20-JUN-2001; 2001WO-US019692.
29-JUN-2001; 2001WO-US021056.
09-JUL-2001; 2001WO-US021735.
18-JUL-2001; 2001US-00908827.
30-JUL-2001; 2001US-00918585.
06-AUG-2001; 2001US-00924419.
13-AUG-2001; 2001US-00929404.
16-AUG-2001; 2001US-00931836.
28-AUG-2001; 2001US-00941992.
29-AUG-2001; 2001WO-US027099.
04-SEP-2001; 2001US-00946374.
15-JAN-2002; 2002US-00052586.
(GETH) GENENTECH INC.
Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
WPI; 2003-332034/31.
N-PSDB; ACA73678.

PT Three hundred and five nucleic acids encoding PRO polypeptides, useful in gene therapy, chromosome identification, tissue typing, and for detecting the presence of tumor in a mammal.
PT Claim 11; Fig 588; 707pp; English.
XX The invention relates to three hundred and five nucleic acids encoding PRO polypeptides (secreted and transmembrane), sequences 80% identical to them, or encoding a PRO polypeptide lacking its associated signal peptide or an extracellular domain of the PRO polypeptide, with or lacking its associated signal peptide. Also included are the encoded PRO proteins, PRO expression vectors, host cells transformed with the vector (used to produce PRO proteins), a chimeric molecule comprising the PRO polypeptide fused to a heterologous amino acid sequence, an anti-PRO antibody, a method for stimulating the release of tumor necrosis factor alpha (TNF-alpha) from human blood (by contacting the blood with PRO1079, PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or PRO4333), a method for stimulating the proliferation or differentiation of chondrocyte cells by contacting the cells with a PRO6023 polypeptide, a method for detecting the presence of tumor in a mammal and an oligonucleotide probe derived from any of the nucleotide sequences cited above. The PRO polypeptide or anti-PRO antibody is useful for preparing a medicament for treating a condition that is responsive to the PRO polypeptide or anti-PRO antibody. The PRO nucleotide sequences are useful as hybridisation probes in chromosome and gene mapping, or in generating antisense RNA and DNA. PRO nucleic acids are also useful in preparing PRO polypeptides, in assays to identify other proteins or molecules involved in a binding reaction, to generate transgenic animals or knockout animals, which in turn are useful in the development and screening of therapeutically useful reagents, for chromosome identification, and tissue typing. The PRO polypeptides and nucleic acid molecules are also useful for detecting the presence of a tumour in a mammal, stimulating the proliferation or differentiation of chondrocyte cells, stimulating the release of tumour necrosis factor-alpha from human blood, in gene therapy, or as molecular weight markers for protein electrophoresis purposes. The anti-PRO antibodies may be used in diagnostic assays for PRO, or for the affinity purification of PRO from recombinant cell culture or natural sources. The present sequence represents a PRO protein Sequence 199 AA;
Qy 3 LVGPSLMSWLTTC 14
Db 184 LTGHSLSMWLVC 195
Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
RESULT 46
ABU67699
ID ABU67699 standard; protein; 199 AA.
XX AC ABU67699;
XX DT 29-MAY-2003 (first entry)
XX DE Human secreted/transmembrane protein (PRO) #294.
XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX OS Homo sapiens.
XX PN US2003036162-A1.
XX PD 20-FEB-2003.
XX PF 12-JUL-2002; 2002US-00194423.
XX PR 26-JUN-1998; 98US-00105413.

PR 16-SEP-1998; 98WO-US019330.
 PR 07-OCT-1998; 98US-00168978.
 PR 07-OCT-1998; 98WO-US021141.
 PR 06-NOV-1998; 98US-00187368.
 PR 01-DEC-1998; 98WO-US025108.
 PR 07-DEC-1998; 98US-00202054.
 PR 03-MAR-1999; 99US-00254311.
 PR 08-MAR-1999; 99WO-US005028.
 PR 14-MAY-1999; 99US-00311832.
 PR 14-MAY-1999; 99WO-US010733.
 PR 14-MAY-1999; 99WO-US010733.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380138.
 PR 25-AUG-1999; 99US-00380139.
 PR 25-AUG-1999; 99US-00380142.
 PR 01-SEP-1999; 99WO-US020111.
 PR 15-SEP-1999; 99WO-US020109.
 PR 18-OCT-1999; 99US-00403297.
 PR 12-NOV-1999; 99US-00423844.
 PR 01-DEC-1999; 99WO-US028301.
 PR 02-DEC-1999; 99WO-US028551.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 24-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US011705.
 PR 22-MAY-2000; 2000WO-US011404.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 22-AUG-2000; 2000US-00644848.
 PR 24-AUG-2000; 2000WO-US021328.
 PR 18-SEP-2000; 2000US-00664610.
 PR 18-SEP-2000; 2000US-00665350.
 PR 08-NOV-2000; 2000US-00709238.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 28-FEB-2001; 2001US-00816744.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 25-MAY-2001; 2001US-00866028.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 30-JUL-2001; 2001US-00918585.
 PR 06-AUG-2001; 2001US-00924419.
 PR 13-AUG-2001; 2001US-00929404.
 PR 16-AUG-2001; 2001US-00931836.
 PR 28-AUG-2001; 2001US-00941992.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 04-SEP-2001; 2001US-00946374.
 PR 15-JAN-2002; 2002US-00052586.
 XX
 PA (GETH) GENENTECH INC.
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2003-332039/31.
 DR N-PSDB; ACA05993.
 XX

PT New secreted and transmembrane PRO polypeptides and nucleic acids, useful
 PT in gene therapy, in chromosome and gene mapping, as chromosome markers,
 PT in tissue typing, and in chromosome identification.
 XX
 PS Claim 11; Fig 588; 706pp; English.
 XX
 CC The invention discloses human nucleic acids encoding secreted and
 CC transmembrane (PRO) polypeptides. Also disclosed is an antibody that
 CC specifically binds to the PRO polypeptide, a method for stimulating the
 CC release of tumour necrosis factor alpha (TNF-alpha) from human blood by
 CC contacting the blood a PRO polypeptide, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells by contacting the
 CC cells with a PRO polypeptide, a method for detecting the presence of a
 CC tumour in a mammal and an oligonucleotide probe derived from any of the
 CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
 CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
 CC preparing PRO polypeptides by recombinant techniques and in gene therapy
 CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
 CC as molecular weight markers for protein electrophoresis purposes, for
 CC chromosome identification, as chromosome markers, as therapeutic agents,
 CC for stimulating the release of TNF-alpha from human blood, for
 CC stimulating the proliferation or differentiation of chondrocytes and
 CC detecting the presence of a tumour. The PRO polypeptides and nucleic
 CC acids may also be used diagnostically for tissue typing. The sequences
 CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention
 XX
 SQ Sequence 199 AA;

 QY 3 LVGPSLSWLTC 14
 DB 184 LTGHLSLSWLVC 195

 RESULT 47
 ABUS0727
 ID ABUS0727 standard; protein; 199 AA.
 XX
 AC ABUS0727;
 DT 23-JUN-2003 (first entry)
 XX
 DE Human PRO protein #294.
 XX
 KW Human; tumour; adrenal; lung; colon; breast; prostate; rectal; cervical;
 KW liver; PRO; gene therapy.
 OS Homo sapiens.
 XX
 PN US2003036137-A1.
 XX
 PD 20-FEB-2003.
 XX
 PF 27-JUN-2002; 2002US-00184640.
 XX
 PR 26-JUN-1998; 98US-00105413.
 PR 16-SEP-1998; 98WO-US019330.
 PR 07-OCT-1998; 98US-00168978.
 PR 07-OCT-1998; 98WO-US021141.
 PR 06-NOV-1998; 98US-00187368.
 PR 01-DEC-1998; 98WO-US025108.
 PR 07-DEC-1998; 98US-00202054.
 PR 03-MAR-1999; 99US-00254311.
 PR 08-MAR-1999; 99WO-US005028.
 PR 14-MAY-1999; 99US-00311832.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380138.
 PR 25-AUG-1999; 99US-00380139.
 PR

Query Match 60.5%; Score 49; DB 6; Length 199;
 Best Local Similarity 75.0%; Pred. No. 6.2;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087988P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088122P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088555P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090441P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match

60.5%; Score 49; DB 6; Length 199;

```
Best Local Similarity 75.0%; Pred. No. 6-2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 LVGPSLMSLTC 14
      | | | | | | |
Db      184 LTGHSLSMLVC 195

RESULT 49
ABR99035
ID ABR99035 standard; protein; 199 AA.
AC ABR99035;
XX
XX
DT 17-SEP-2003 (first entry)
DE XX
DE Human secreted polypeptide PRO19669, SEQ ID NO:588.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
XX US2003040064-A1.
XX
PD 27-FEB-2003.
XX
XX 26-JUN-2002; 2002US-001833008.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 11-DEC-1997; 97US-0066772P.
PR 12-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 15-APR-1998; 98US-0081195P.
PR 21-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 22-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 12-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
```


PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088333P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088718P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091488P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.

PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.

Query Match 60.5%; Score 49; DB 6; Length 199;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGSPSLMSWLTTC 14
DB 194 LTGHSLMSWLVLC 195

Search completed: May 12, 2006, 10:36:06
Job time : 130.809 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 16.9915 Seconds
(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-67

Perfect score: 81

Sequence: 1 CELVGPSIMSWLTC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR_80:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	56.8	365	2 T30343	irp3 protein - Yer
2	46	56.8	365	2 T17441	hypothetical prote
3	46	56.8	366	2 A10232	yersiniabactin bio
4	45	55.6	907	2 AB1885	hypothetical prote
5	44	54.3	1132	2 A35098	MHC class III hist
6	43	53.1	224	2 T35918	hypothetical prote
7	42.5	52.5	261	2 C55581	nasD protein - Kle
8	42	51.9	538	2 T00424	probable Na+/Ca2+
9	42	51.9	686	2 S33475	hypothetical prote
10	42	51.9	786	2 G88065	protein T16A1.2 [i
11	42	51.9	869	2 S49844	probable membrane
12	41	50.6	263	1 EQ80A	enkephalin e-12 pr
13	41	50.6	290	2 AB7380	conserved hypothet
14	41	50.6	361	2 H71631	queuine tRNA-ribos
15	41	50.6	389	2 S26414	naringenin-chalcon
16	41	50.6	703	2 T03888	Na+/Ca2+, K+-exchan
17	41	50.6	760	2 T34414	hypothetical prote
18	40	49.4	84	2 AB2101	hypothetical prote
19	40	49.4	372	2 T02735	3-methyl-2-oxobuta
20	40	49.4	410	1 DEPSXA	2-oxoisovalerate d
21	40	49.4	410	2 C83365	naphthalene dioxyg
22	40	49.4	459	2 T31134	hypothetical prote
23	40	49.4	651	2 T03889	Na+/Ca2+, K+-exchan
24	40	49.4	1068	2 T04112	pol protein homolo
25	39	48.1	266	2 S54440	hem-in-specific ATP
26	39	48.1	306	2 C85224	probable LTR retro
27	39	48.1	306	2 T04759	hypothetical prote
28	39	48.1	465	2 T48374	UDPG glucosyltrans
29	39	48.1	735	2 S46830	urea transport pro

30	38	46.9	112	2 S75152	hypothetical prote
31	38	46.9	219	2 A46097	GPI-anchor biosynt
32	38	46.9	247	2 C90848	hypothetical prote
33	38	46.9	247	2 B85706	hypothetical prote
34	38	46.9	247	2 B64873	probable membrane
35	38	46.9	267	1 EQHUA	enkephalin precurs
36	38	46.9	286	2 C95847	hypothetical trans
37	38	46.9	342	2 T09355	hypothetical prote
38	38	46.9	357	2 T12379	NAHD2 dehydrogenas
39	38	46.9	437	1 B29336	ubiquinol-cytochro
40	38	46.9	441	2 B69004	hypothetical prote
41	38	46.9	455	2 AG0308	isochorismate synt
42	38	46.9	472	2 A13367	multidrug resistanc
43	38	46.9	493	2 T48630	high affinity nitr
44	38	46.9	499	2 H83254	probable NFS trans
45	38	46.9	541	1 A70022	multidrug-efflux t
46	38	46.9	594	2 T22072	hypothetical prote
47	38	46.9	644	2 T22067	hypothetical prote
48	38	46.9	718	2 H75485	hypothetical prote
49	38	46.9	846	2 T04970	probable potassium
50	38	46.9	941	2 B96553	hypothetical prote
51	38	46.9	1149	2 I38006	M130 antigen precu
52	38	46.9	1151	2 I38004	M130 antigen precu
53	38	46.9	1156	2 I38005	M130 antigen precu
54	37.5	46.3	571	2 A75499	DNA primase - Dein
55	37	45.7	248	2 H83111	hypothetical prote
56	37	45.7	260	2 S72748	B1177_F3_136 prote
57	37	45.7	266	2 T12073	ATP-binding protei
58	37	45.7	266	2 AC0035	hem-in transport sy
59	37	45.7	297	2 AC0615	conserved hypothet
60	37	45.7	297	2 B90754	hypothetical prote
61	37	45.7	297	2 F64831	hypothetical prote
62	37	45.7	297	2 H85617	hypothetical prote
63	37	45.7	333	2 D95968	probable amino aci
64	37	45.7	371	2 D96501	hypothetical prote
65	37	45.7	378	2 D83381	hypothetical prote
66	37	45.7	387	2 T40309	cytochrome c oxida
67	37	45.7	445	2 T01591	hypothetical prote
68	37	45.7	602	2 T48697	hypothetical prote
69	37	45.7	868	2 T25716	hypothetical prote
70	37	45.7	1121	2 T25715	hypothetical prote
71	37	45.7	1240	2 T04833	hypothetical prote
72	37	45.7	1309	2 H96650	protein T3P18.3 [i
73	37	45.7	1415	2 C86438	hypothetical prote
74	37	45.7	1796	2 AC1895	serine/threonine k
75	37	45.7	3856	2 T51174	ataxia-telangiecta
76	36.5	45.1	229	2 H72110	ribulose-phosphate
77	36.5	45.1	229	2 A86514	ribulose-P epimera
78	36.5	45.1	261	2 A60091	per region circadi
79	36.5	45.1	1572	2 T00027	brain-specific ang
80	36	44.4	166	2 F96786	protein F10A5.11 [
81	36	44.4	199	2 T45543	hypothetical prote
82	36	44.4	204	2 T28939	hypothetical prote
83	36	44.4	233	2 S74430	hypothetical prote
84	36	44.4	247	2 T45502	hypothetical prote
85	36	44.4	304	2 AD1029	probable membrane
86	36	44.4	313	1 S75303	hypothetical prote
87	36	44.4	315	2 T11388	alpha-antigen - My
88	36	44.4	330	2 S32773	hypothetical prote
89	36	44.4	335	2 T04029	chloromuconate cyc
90	36	44.4	370	1 B35255	ubiquinol-cytochro
91	36	44.4	382	2 T11219	hypothetical prote
92	36	44.4	446	2 DB1058	probable glucarat
93	36	44.4	455	1 H69752	multidrug-efflux t
94	36	44.4	477	2 B75409	hypothetical prote
95	36	44.4	498	2 T04640	xylan 1,4-beta-xy
96	36	44.4	517	2 A49776	conserved hypothet
97	36	44.4	524	2 D82220	hypothetical prote
98	36	44.4	538	2 H96008	conserved hypothet
99	36	44.4	558	2 F82236	hypothetical prote
100	36	44.4	567	2 D70769	hypothetical prote
101	36	44.4	575	2 AG2566	hypothetical prote
102	36	44.4	576	2 S33327	catalase (EC 1.11.

103	36	44.4	588	2	A82573	hypothetical prote	176	35	43.2	570	2	S42708	proline transport
104	36	44.4	592	2	T51712	threonine ammonia-	177	35	43.2	570	2	S04547	proline transport
105	36	44.4	597	2	S85673	GABA transport pro	178	35	43.2	576	2	T12005	NADH2 dehydrogenas
106	36	44.4	676	1	WMBX6	UL6 protein - huma	179	35	43.2	576	2	AF2361	flavoprotein [limpo
107	36	44.4	776	2	T09938	hypothetical prote	180	35	43.2	580	2	T09808	NADH2 dehydrogenas
108	36	44.4	1014	2	P96501	hypothetical prote	181	35	43.2	591	1	C8HUB	complement C8 beta
109	36	44.4	1158	2	S86327	protein F18014.19	182	35	43.2	635	2	A87433	hypothetical prote
110	36	44.4	1257	2	S28764	neurocan precursor	183	35	43.2	659	2	T02838	probable membrane
111	36	44.4	1267	2	T21340	hypothetical prote	184	35	43.2	676	2	D96615	unknown protein F1
112	36	44.4	1404	1	A48196	hypothetical prote	185	35	43.2	694	2	S67296	unknown protein F2
113	36	44.4	1677	2	T18344	protein-tyrosine k	186	35	43.2	699	2	D96767	probable membrane
114	36	44.4	1680	2	T01367	P-glycoprotein E -	187	35	43.2	716	2	S26759	peg protein - frui
115	36	44.4	1972	2	S68176	hypothetical prote	188	35	43.2	762	2	H87466	beta-D-glucosidase
116	36	44.4	2110	2	B44110	TOG protein - huma	189	35	43.2	819	2	T08745	probable RNA helic
117	36	44.4	2111	2	A70688	mycocerosate synth	190	35	43.2	819	2	T05744	hypothetical prote
118	36	44.4	2118	2	T13612	mycocerosate synth	191	35	43.2	838	2	G84599	hypothetical prote
119	36	44.4	2264	1	GNVVTB	hypothetical prote	192	35	43.2	879	2	S23006	probable retroelem
120	36	44.4	3033	1	J01303	genome polypotein	193	35	43.2	886	1	RGBYL3	shed acute-phase a
121	35.5	43.8	402	2	A85595	probable DEOR-type	194	35	43.2	897	2	B69202	regulatory protein
122	35.5	43.8	402	2	E30744	probable DEOR-type	195	35	43.2	929	2	S75098	valine-tRNA ligase
123	35.5	43.8	519	2	S45345	TUPI-like enhancer	196	35	43.2	947	2	T01238	hypothetical prote
124	35.5	43.8	766	2	S45344	TUPI like enhancer	197	35	43.2	993	2	G96680	hypothetical prote
125	35.5	43.8	1015	2	S68141	nuclear protein HI	198	35	43.2	1015	2	B84458	probable retroelem
126	35.5	43.8	1017	2	I37445	HIRA protein - hum	199	35	43.2	1104	2	T01811	hypothetical prote
127	35	43.2	135	2	AC2305	hypothetical prote	200	35	43.2	1135	2	T30561	Scythe protein - A
128	35	43.2	148	2	G87639	conserved hypothet	201	35	43.2	1294	2	S58307	hypothetical 149.2
129	35	43.2	156	1	GNVQ12	genome-linked prot	202	35	43.2	1307	2	T21283	hypothetical prote
130	35	43.2	156	1	GNVQ12	genome-linked prot	203	35	43.2	1318	2	T05745	hypothetical prote
131	35	43.2	156	1	GNVQVA	genome-linked prot	204	35	43.2	1439	2	T02087	gag/pol polyprotei
132	35	43.2	161	2	T22196	hypothetical prote	205	35	43.2	1489	2	G71406	probable retroviri
133	35	43.2	190	2	C89966	hypothetical prote	206	35	43.2	1505	2	S28079	DNA-directed DNA p
134	35	43.2	191	2	S41337	hypothetical prote	207	35	43.2	1545	1	S71841	multidrug resistan
135	35	43.2	212	2	AC2326	hypothetical prote	208	35	43.2	1908	2	T42707	hypothetical prote
136	35	43.2	220	2	S54427	HUA-DR beta-chain	209	35	43.2	1952	2	T48814	hypothetical prote
137	35	43.2	233	2	C30427	hypothetical prote	210	34.5	42.6	235	2	T30656	hypothetical prote
138	35	43.2	246	2	I59092	HUA-DR beta-chain	211	34.5	42.6	236	1	TVMSA1	transforming prote
139	35	43.2	258	1	DHMS	Ig delta chain C r	212	34.5	42.6	236	1	I53744	transforming prote
140	35	43.2	262	1	EQXL	enkephalin precurs	213	34.5	42.6	239	1	TVHUA1	transforming prote
141	35	43.2	265	2	A70640	hypothetical prote	214	34.5	42.6	335	2	S58997	NADH2 dehydrogenas
142	35	43.2	266	2	B28031	MHC class II hista	215	34.5	42.6	402	1	B64822	probable membrane
143	35	43.2	301	1	DHSM	Ig delta chain C r	216	34.5	42.6	475	2	H84137	hypothetical prote
144	35	43.2	302	2	S75227	hypothetical prote	217	34.5	42.6	582	2	S08686	finger protein zpp
145	35	43.2	308	2	AB1957	hypothetical prote	218	34.5	42.6	671	2	JE0288	kruepel-type zinc
146	35	43.2	308	2	T38947	hypothetical homoc	219	34.5	42.6	999	1	S27756	receptor-like prot
147	35	43.2	326	2	C71426	hypothetical prote	220	34.5	42.6	1097	1	RNBXL3	DNA-directed DNA p
148	35	43.2	329	2	T11350	hypothetical prote	221	34.5	42.6	1299	2	T43251	furin (EC 3.4.21.7
149	35	43.2	334	2	AH3282	primosomal prote	222	34	42.0	62	2	T13997	aquaporin - common
150	35	43.2	335	2	A86454	hypothetical prote	223	34	42.0	129	2	A82227	conserved hypothet
151	35	43.2	339	2	S42595	hypothetical prote	224	34	42.0	138	2	S34735	beta-galactosidase
152	35	43.2	344	2	A51626	nickel resistance	225	34	42.0	143	2	T16864	hypothetical prote
153	35	43.2	348	2	A52957	zinc-binding dehyd	226	34	42.0	144	2	AC1907	hypothetical prote
154	35	43.2	348	2	D98326	probable zinc-bind	227	34	42.0	147	2	JC7938	type II antifreeze
155	35	43.2	357	2	T12448	NADH2 dehydrogenas	228	34	42.0	160	2	I61895	eosinophil cationi
156	35	43.2	364	2	A30521	myeloid cell surfa	229	34	42.0	172	2	AD3606	molybdopterin bios
157	35	43.2	379	2	I48133	ubiquinol-cytochro	230	34	42.0	176	1	JQ1636	major noncaped pr
158	35	43.2	379	2	I48132	ubiquinol-cytochro	231	34	42.0	204	2	P72459	hypothetical prote
159	35	43.2	379	2	I48134	ubiquinol-cytochro	232	34	42.0	220	2	A25925	class II histocomp
160	35	43.2	382	2	T11138	ubiquinol-cytochro	233	34	42.0	240	2	S53403	probable membrane
161	35	43.2	384	2	S62758	ubiquinol-cytochro	234	34	42.0	244	2	B81690	probable sodium-tr
162	35	43.2	407	2	T00999	hypothetical prote	235	34	42.0	247	2	AC0652	probable membrane
163	35	43.2	414	2	S75052	hypothetical prote	236	34	42.0	250	2	S47037	tonoplast intrinsi
164	35	43.2	417	2	D82111	uracil permease VC	237	34	42.0	250	2	S52004	gamma-Tip protein
165	35	43.2	420	2	A64132	amphotropic murine	238	34	42.0	250	2	D82258	cy52 protein VC096
166	35	43.2	421	2	A56550	Krox-20 - African	239	34	42.0	253	2	T12439	tonoplast intrinsi
167	35	43.2	427	1	T49031	acid phosphatase (240	34	42.0	259	2	T52619	TINY-like protein
168	35	43.2	443	2	H84414	hypothetical prote	241	34	42.0	265	2	B39797	MHC class II hista
169	35	43.2	447	2	T49990	hypothetical prote	242	34	42.0	266	1	HLHU4D	MHC class II hista
170	35	43.2	447	2	P97393	hypothetical prote	243	34	42.0	266	2	A28031	MHC class II hista
171	35	43.2	477	2	B83476	probable transport	244	34	42.0	269	1	EQRTA	enkephalin A precu
172	35	43.2	501	2	E96758	hypothetical prote	245	34	42.0	269	2	B35678	enkephalin precurs
173	35	43.2	532	2	A30766	beta-N-acetylhexos	246	34	42.0	276	2	C82420	lipase-related pro
174	35	43.2	536	2	T41652	hypothetical prote	247	34	42.0	278	2	S49623	hydroxynuosoporen
175	35	43.2	536	2	A82133	ferric aerobactin	248	34	42.0	279	2	T50748	hydroxynuosoporen

249	34	42.0	291	2	T40277	probable membrane	322	33.5	41.4	472	2	H96607	unknown protein F2
250	34	42.0	304	2	AG3035	hypothetical prote	323	33.5	41.4	548	2	B87596	xyloisidase/arabino
251	34	42.0	304	2	D98250	probable sugar ABC	324	33.5	41.4	580	2	T33742	hypothetical prote
252	34	42.0	317	2	S45054	GTP-binding regula	325	33.5	41.4	619	2	S67067	probable membrane
253	34	42.0	317	2	S38398	GTP-binding regula	326	33.5	41.4	1038	1	JC5757	DNA-directed DNA p
254	34	42.0	317	2	A33928	GTP-binding protei	327	33.5	41.4	1038	2	T18222	DNA polymerase del
255	34	42.0	317	2	B33928	GTP-binding protei	328	33.5	41.4	1247	2	T45743	hypothetical prote
256	34	42.0	317	2	A36986	activated protein	329	33.5	41.4	1496	2	AH0447	insecticidal toxin
257	34	42.0	324	2	T01316	epoxide hydrolase	330	33.5	41.4	2061	2	T13751	transcription fact
258	34	42.0	329	2	T07000	chitinase [EC 3.2.	331	33.5	41.4	2318	2	S45306	notch 3 protein -
259	34	42.0	336	2	A46478	major surface anti	332	33.5	41.4	4574	2	G02520	plectin - human
260	34	42.0	336	2	A30527	major surface anti	333	33.5	41.4	4684	2	A59404	ig heavy chain v r
261	34	42.0	343	2	S62704	ubiquinol-cytochro	334	33.5	41.4	55	2	S20771	ATP synthase F0 ch
262	34	42.0	357	2	T11703	NADH2 dehydrogenas	335	33	40.7	88	2	E90618	hypothetical prote
263	34	42.0	360	2	A80058	phospho-N-acetylmu	336	33	40.7	103	2	S59846	hypothetical prote
264	34	42.0	370	1	B43673	chloromuconate cyc	337	33	40.7	106	2	D69808	hypothetical prote
265	34	42.0	377	2	F95982	probable transport	338	33	40.7	112	2	F75079	hypothetical prote
266	34	42.0	393	2	D95894	probable hydrolase	339	33	40.7	128	2	T22276	hypothetical prote
267	34	42.0	394	2	S75996	hypothetical prote	340	33	40.7	138	2	E65019	hypothetical prote
268	34	42.0	396	2	T23619	hypothetical prote	341	33	40.7	142	2	A19334	hypothetical prote
269	34	42.0	430	2	B70959	probable dihydroor	342	33	40.7	149	2	S74366	hypothetical prote
270	34	42.0	435	1	W2BES0	gene 50 protein -	343	33	40.7	156	2	AG2741	conserved hypotet
271	34	42.0	449	2	F81954	ubiquinol-cytochro	344	33	40.7	184	2	AE1834	hypothetical prote
272	34	42.0	449	2	E81011	ubiquinol-cytochro	345	33	40.7	166	2	I39546	hypothetical prote
273	34	42.0	449	2	C81817	probable integral	346	33	40.7	172	2	F97522	hypothetical prote
274	34	42.0	477	2	H83588	probable MFS trans	347	33	40.7	185	2	I40478	hypothetical prote
275	34	42.0	478	2	E86314	F2H15.15 protein -	348	33	40.7	194	2	H85097	hypothetical prote
276	34	42.0	483	2	T10532	hypothetical prote	349	33	40.7	202	2	G85439	cold acclimation p
277	34	42.0	486	2	F86654	transporter ycdH [350	33	40.7	231	2	S60769	DNA repair/recombi
278	34	42.0	503	2	T05347	hypothetical prote	351	33	40.7	232	2	AF3380	hypothetical prote
279	34	42.0	511	1	E70391	major facilitator	352	33	40.7	232	2	C70699	probable paba prot
280	34	42.0	532	1	B69821	multidrug resistan	353	33	40.7	238	2	T28337	probable metallopr
281	34	42.0	535	2	B84239	hypothetical prote	354	33	40.7	240	1	JJAGTT	adenylate isopente
282	34	42.0	561	2	AC0605	probable membrane	355	33	40.7	240	2	T12191	NADH2 dehydrogenas
283	34	42.0	561	2	G90744	probable transport	356	33	40.7	240	2	AB1871	hypothetical prote
284	34	42.0	561	2	C85595	probable membrane	357	33	40.7	246	2	D83540	probable paba prot
285	34	42.0	561	2	G84822	TrkA, Potassium ch	358	33	40.7	247	2	B84962	probable metallopr
286	34	42.0	562	2	A10161	hypothetical prote	359	33	40.7	249	2	A70840	adenylate isopente
287	34	42.0	566	2	T15226	6Fe-6S prismane cl	360	33	40.7	252	2	T01947	NADH2 dehydrogenas
288	34	42.0	567	2	C97321	hypothetical prote	361	33	40.7	252	2	T01947	hypothetical prote
289	34	42.0	574	2	AH2292	hypothetical prote	362	33	40.7	253	2	T14518	tonoplast intrinsi
290	34	42.0	618	2	A10171	probable exported	363	33	40.7	256	2	JQ1106	tonoplast intrinsi
291	34	42.0	635	2	S22999	trAG protein - Esc	364	33	40.7	264	2	AH0759	panepanediol diffus
292	34	42.0	637	2	T08530	trAG protein - Esc	365	33	40.7	264	2	T10253	membrane protein M
293	34	42.0	637	2	S22992	trAG protein - Esc	366	33	40.7	269	2	H70706	probable OXIPORSDU
294	34	42.0	643	2	G75599	probable drug tran	367	33	40.7	275	2	T10251	membrane protein M
295	34	42.0	687	2	AG2385	hypothetical prote	368	33	40.7	279	2	T10251	hydroxyneurosporen
296	34	42.0	758	2	D87369	beta-D-glucosidase	369	33	40.7	281	2	S04405	probable membrane
297	34	42.0	766	2	T38120	hypothetical prote	370	33	40.7	282	2	S61663	hypothetical prote
298	34	42.0	803	2	T39530	hypothetical prote	371	33	40.7	284	2	B82955	conserved hypotet
299	34	42.0	884	2	S77031	hypothetical prote	372	33	40.7	289	2	AB0262	hypothetical prote
300	34	42.0	940	2	T01854	hypothetical prote	373	33	40.7	290	2	T23268	hypothetical prote
301	34	42.0	974	2	A72012	metalloproteinase,	374	33	40.7	306	2	T26639	trans-activating t
302	34	42.0	974	2	B86613	zinc metalloprotei	375	33	40.7	308	1	QOLJX1	NADH2 dehydrogenas
303	34	42.0	975	2	T29908	hypothetical prote	376	33	40.7	311	2	T11362	hypothetical prote
304	34	42.0	985	2	I51672	receptor tyrosine	377	33	40.7	320	2	A12305	transcription regu
305	34	42.0	1008	2	D84434	probable receptor-	378	33	40.7	322	2	AG3201	transcription regu
306	34	42.0	1025	2	A82516	type I restriction	379	33	40.7	324	2	G82410	hypothetical prote
307	34	42.0	1064	2	A41542	adenylate cyclase	380	33	40.7	325	2	D70728	hypothetical prote
308	34	42.0	1179	2	T05673	hypothetical prote	381	33	40.7	326	2	AH2612	NADH-ubiquinone ox
309	34	42.0	1404	2	F86470	probable retroelem	382	33	40.7	328	2	I47161	ig gamma 3 chain c
310	34	42.0	1413	2	G86301	probable retroelem	383	33	40.7	329	2	T40158	ig gamma 1 chain c
311	34	42.0	1435	1	BVBYL1	guanine nucleotide	384	33	40.7	329	2	T40158	hypothetical prote
312	34	42.0	1468	2	T05672	hypothetical prote	385	33	40.7	340	2	S49742	hypothetical prote
313	34	42.0	1567	2	T03730	antigen containing	386	33	40.7	342	2	E84905	probable transport
314	34	42.0	1777	2	AC2088	serine/threonine k	387	33	40.7	342	2	A90894	probable transport
315	34	42.0	2476	2	T34022	zonadhesin - pig	388	33	40.7	343	2	A35639	probable sugar tra
316	34	42.0	6658	2	T13931	projectin - fruit	389	33	40.7	345	2	G86372	gene alx3 protein
317	33.5	41.4	236	2	JC7383	B-cell lymphoma 2	390	33	40.7	346	2	I48185	trans-1,2-dihydrob
318	33.5	41.4	245	2	C86418	hypothetical prote	391	33	40.7	349	2	AD3650	fatty-acid desatur
319	33.5	41.4	343	2	B97256	D-alanine-D-alanin	392	33	40.7	352	2	B69901	probable NADH-ubiq
320	33.5	41.4	353	2	AF2053	hypothetical prote	393	33	40.7	354	2	F97394	
321	33.5	41.4	399	2	AC1601	hypothetical prote	394	33	40.7	354	2	F97394	

395	33	40.7	355	2	C64644	integrase/recombin	468	33	40.7	578	2	T48795	origin recognition
396	33	40.7	361	2	A97837	queuine tRNA-ribos	469	33	40.7	582	2	H70588	probable pSOS prot
397	33	40.7	362	2	G69589	3-dehydroquinat s	470	33	40.7	592	2	T09867	hypothetical prote
398	33	40.7	367	2	T06706	hypothetical prote	471	33	40.7	612	2	T37816	NADH2 dehydrogenas
399	33	40.7	370	2	T44667	chloromuconate cyc	472	33	40.7	614	2	A41757	betaine transport
400	33	40.7	370	2	T44618	chloromuconate cyc	473	33	40.7	622	2	AG0952	membrane transport
401	33	40.7	370	2	JQ0176	chloromuconate cyc	474	33	40.7	622	2	A98215	low affinity potas
402	33	40.7	371	2	B86489	protein T32E20.25	475	33	40.7	622	2	A43345	K(+)-uptake protei
403	33	40.7	379	2	T48135	ubiquinol-cytochro	476	33	40.7	622	2	B86061	low affinity potas
404	33	40.7	379	2	T48180	ubiquinol-cytochro	477	33	40.7	624	2	T49366	myocyte-specific e
405	33	40.7	380	2	A71390	ubiquinol-cytochro	478	33	40.7	635	2	AE2086	hypothetical prote
406	33	40.7	382	2	D58930	ubiquinol-cytochro	479	33	40.7	653	2	A49722	endoglin precursor
407	33	40.7	399	2	T51472	hypothetical prote	480	33	40.7	655	1	A46688	hepatocyte growth
408	33	40.7	400	2	A41579	beta-3-adrenergic	481	33	40.7	658	2	S50831	endoglin precursor
409	33	40.7	400	2	A53281	beta 3-adrenergic	482	33	40.7	660	2	G82672	Arp sulphyrase, 1
410	33	40.7	400	2	T49810	hypothetical prote	483	33	40.7	662	2	T44219	hypothetical prote
411	33	40.7	402	2	A70882	probable pPS prote	484	33	40.7	662	2	T44034	hypothetical prote
412	33	40.7	405	2	T44249	transport protein	485	33	40.7	687	2	B70515	hypothetical prote
413	33	40.7	407	2	T36404	probable monooxyge	486	33	40.7	688	2	T05353	hypothetical prote
414	33	40.7	409	2	T08928	sucrose cleavage p	487	33	40.7	714	2	T14080	hypothetical prote
415	33	40.7	412	2	G87522	hypothetical prote	488	33	40.7	756	2	T34797	probable Arp bindi
416	33	40.7	415	2	B97512	probable cytochrom	489	33	40.7	764	2	T01493	probable vacuolar
417	33	40.7	415	2	AC2731	cytochrome P450 hy	490	33	40.7	780	2	H84685	probable ABC trans
418	33	40.7	416	2	A37877	hemorrhagic protei	491	33	40.7	795	2	T34673	probable SecDF pro
419	33	40.7	418	2	B33686	hypothetical prote	492	33	40.7	856	2	T22575	hypothetical prote
420	33	40.7	422	2	G83503	probable MFS trans	493	33	40.7	873	2	T12535	hypothetical prote
421	33	40.7	422	2	A71058	probable cytosine	494	33	40.7	894	2	B96557	probable receptor
422	33	40.7	427	2	S30558	ribose-phosphate d	495	33	40.7	899	2	T07062	probable lipoxigen
423	33	40.7	427	2	C86469	protein F12K21.14	496	33	40.7	911	2	AD2271	cation-transportin
424	33	40.7	435	2	T29593	hypothetical prote	497	33	40.7	992	2	T39817	hypothetical prote
425	33	40.7	436	1	A94558	diphosphate-P450 fa	498	33	40.7	1007	2	T24643	hypothetical prote
426	33	40.7	444	2	F87366	probable dipeptida	499	33	40.7	1070	1	A54600	1-phosphatidylinos
427	33	40.7	448	2	C83347	probable dipeptida	500	33	40.7	1076	2	T30842	serine-repeat anti
428	33	40.7	449	2	F83328	probable sodium/al	501	33	40.7	1088	1	PFRTGA	guanylate cyclase
429	33	40.7	451	2	A55909	transforming prote	502	33	40.7	1100	2	T42260	WD-repeat protein
430	33	40.7	456	1	I40516	spaf protein - Bac	503	33	40.7	1121	2	GC7329	probable ABC trans
431	33	40.7	456	2	B83750	gluconate permease	504	33	40.7	1194	1	J70837	unknown protein [i
432	33	40.7	456	2	A83264	hypothetical prote	505	33	40.7	1311	2	G86471	hypothetical prote
433	33	40.7	462	2	AD1284	menaquinone-specif	506	33	40.7	1385	2	T13415	hypothetical prote
434	33	40.7	462	2	AG1655	hypothetical prote	507	33	40.7	1415	2	T08945	probable LTR retro
435	33	40.7	464	2	T16889	conserved hypotet	508	33	40.7	1433	2	A71444	probable retroelem
436	33	40.7	466	2	H87419	HBG-G2 (HFK-2) pro	509	33	40.7	1454	2	E84535	probable retroelem
437	33	40.7	469	2	I37451	transcription fact	510	33	40.7	1461	2	E84589	probable retroelem
438	33	40.7	476	2	A54743	NADH oxidase nox -	511	33	40.7	1474	2	D88550	protein ZC84.6 [im
439	33	40.7	479	2	S73770	brain factor 1 pro	512	33	40.7	1502	2	D84587	probable myosin he
440	33	40.7	480	2	JH0672	protein T23E18.18	513	33	40.7	1522	2	C96578	hypothetical prote
441	33	40.7	482	2	A96790	multidrug-efflux t	514	33	40.7	1541	1	S71839	canalicul multidi
442	33	40.7	489	2	A11376	multidrug-efflux t	515	33	40.7	1697	2	T00079	hypothetical prote
443	33	40.7	489	2	A11639	hypothetical prote	516	33	40.7	2049	2	T29227	genome polyprotein
444	33	40.7	489	2	G71245	hypothetical prote	517	33	40.7	2210	1	RRXPTV	hypothetical prote
445	33	40.7	491	2	T48989	drug-export protei	518	33	40.7	2824	2	T22759	hypothetical prote
446	33	40.7	493	2	AD1398	drug-export protei	519	33	40.7	2844	2	S28291	genome polyprotein
447	33	40.7	493	2	AG1773	probable pectinest	520	33	40.7	3033	1	GNWVJ8	dynamin heavy chain
448	33	40.7	496	2	T00978	beta3-glycosyltran	521	33	40.7	4344	1	A53489	protein T21E12.4 [
449	33	40.7	498	2	JC8008	hypothetical prote	522	33	40.7	4464	2	D87755	plectin - rat
450	33	40.7	502	2	C86263	conserved membrane	523	33	40.7	4687	1	A39638	dynamin heavy chain
451	33	40.7	506	2	B87102	probable monooacch	524	33	40.7	4725	1	A44357	zonadhesin - mouse
452	33	40.7	513	2	T14864	sodium/alanine sym	525	33	40.7	5376	2	T42215	serum amyloid A3 p
453	33	40.7	516	2	D82279	kup protein - Esch	526	32.5	40.1	75	2	AC2441	hypothetical prote
454	33	40.7	519	2	D65178	hypothetical prote	527	32.5	40.1	122	4	JN0029	hypothetical prote
455	33	40.7	525	2	T47409	interleukin 2 rece	528	32.5	40.1	140	2	D96732	hypothetical prote
456	33	40.7	537	2	B46535	metalloproteinase	529	32.5	40.1	174	2	T38498	hypothetical prote
457	33	40.7	549	2	S48169	hypothetical prote	530	32.5	40.1	359	2	T30382	alcohol dehydrogen
458	33	40.7	549	2	B86337	NADH2 dehydrogenas	531	32.5	40.1	375	1	B49107	hypothetical prote
459	33	40.7	552	2	T11161	NADH2 dehydrogenas	532	32.5	40.1	401	2	AG1238	hypothetical prote
460	33	40.7	567	2	C85643	hypothetical prote	533	32.5	40.1	403	2	B46165	envelope surface g
461	33	40.7	567	2	H90782	hypothetical prote	534	32.5	40.1	442	2	D84224	heme synthase limo
462	33	40.7	572	2	T11478	NADH2 dehydrogenas	535	32.5	40.1	528	2	T43919	yfub protein limpo
463	33	40.7	572	2	T16865	hypothetical prote	536	32.5	40.1	654	2	AD2932	hypothetical prote
464	33	40.7	573	2	H25797	NADH2 dehydrogenas	537	32.5	40.1	654	2	D98350	transcription fact
465	33	40.7	574	2	S01186	hemagglutinin-neur	538	32.5	40.1	869	2	A55384	penicillin-binding
466	33	40.7	576	1	HNNZS	alkaline phosphata	539	32.5	40.1	873	2	B75514	preprotein translo
467	33	40.7	576	2	H82777		540	32.5	40.1	968	2	G81743	

541	32.5	40.1	969	2	G71482	probable protein t	614	32	39.5	308	2	F95903	conserved hypothet
542	32.5	40.1	970	2	A72028	preprotein translo	615	32	39.5	311	2	T32118	hypothetical prote
543	32.5	40.1	970	2	G86595	protein translocas	616	32	39.5	321	2	T11274	NADH2 dehydrogenas
544	32.5	40.1	1986	2	G86595	probable polyketid	617	32	39.5	331	2	T35100	probable secreted
545	32	39.5	53	2	T31169	hypothetical prote	618	32	39.5	335	2	T48319	hypothetical prote
546	32	39.5	61	2	T47055	hypothetical prote	619	32	39.5	338	2	A44132	D-aspartate oxidas
547	32	39.5	66	2	S09855	hypothetical prote	620	32	39.5	339	2	H83252	glycosyltransferas
548	32	39.5	87	2	S15699	hypothetical prote	621	32	39.5	340	2	H91084	probable glucarate
549	32	39.5	93	2	T17447	hypothetical prote	622	32	39.5	341	2	S57595	mitochondrial repl
550	32	39.5	93	2	AC0232	probable transcrip	623	32	39.5	342	2	JC7110	brain-specific mem
551	32	39.5	102	2	S69752	hypothetical prote	624	32	39.5	343	2	A41748	lumican precursor
552	32	39.5	102	2	T49651	hypothetical prote	625	32	39.5	351	1	VCVETC	coat protein - tur
553	32	39.5	113	2	S26277	T cell receptor be	626	32	39.5	354	2	G71465	hypothetical prote
554	32	39.5	113	2	T38318	T-cell receptor be	627	32	39.5	354	2	A81728	conserved hypothet
555	32	39.5	113	2	S26276	hypothetical prote	628	32	39.5	360	2	S00834	int-1-like protein
556	32	39.5	115	2	C83574	hypothetical prote	629	32	39.5	360	2	B36470	Wnt-2 protein - mo
557	32	39.5	120	2	S06731	ig kappa chain pre	630	32	39.5	361	2	T40857	hypothetical prote
558	32	39.5	121	2	S29002	trypsin inhibitor	631	32	39.5	362	2	E96587	hypothetical prote
559	32	39.5	125	2	F81197	holo-(acyl-carrier	632	32	39.5	364	2	AF0421	probable sugar ABC
560	32	39.5	132	2	C97738	hypothetical prote	633	32	39.5	365	2	A85930	probable glucarate
561	32	39.5	134	2	B75488	hypothetical prote	634	32	39.5	367	2	S59146	ubiquinol-cytochro
562	32	39.5	136	2	G71601	hypothetical prote	635	32	39.5	368	2	T04532	probable (S)-2-hyd
563	32	39.5	143	2	H72610	hypothetical prote	636	32	39.5	368	2	G85206	glycolate oxidase-
564	32	39.5	144	2	S38391	T-cell receptor be	637	32	39.5	371	2	C87446	queuine tRNA ribos
565	32	39.5	147	2	D69164	hypothetical prote	638	32	39.5	371	2	G87644	hypothetical prote
566	32	39.5	150	1	W6WL6	E6 protein - human	639	32	39.5	379	2	AG0585	citrate utilizatio
567	32	39.5	151	2	A97008	transcriptional regu	640	32	39.5	383	1	DHHU	ig delta chain C r
568	32	39.5	157	2	T34552	hypothetical prote	641	32	39.5	384	2	T11931	ubiquinol-cytochro
569	32	39.5	161	2	G81544	conserved hypothet	642	32	39.5	385	2	F70604	probable fadE30 pr
570	32	39.5	162	2	C98011	conserved hypothet	643	32	39.5	387	2	A10511	probable metabolit
571	32	39.5	172	2	T34785	hypothetical prote	644	32	39.5	387	2	E97028	probable amidohydr
572	32	39.5	172	2	A40462	probable exported	645	32	39.5	392	2	A13141	Xaa-pro dipeptidas
573	32	39.5	182	2	D65005	hypothetical prote	646	32	39.5	392	2	C98146	Xaa-pro dipeptidas
574	32	39.5	182	2	C85874	probable transport	647	32	39.5	395	2	B96610	hypothetical prote
575	32	39.5	182	2	B91030	probable transpor	648	32	39.5	397	2	T35713	probable oxidoredu
576	32	39.5	182	2	G70687	hypothetical prote	649	32	39.5	399	2	F83796	multidrug-efflux t
577	32	39.5	196	2	S45553	spore maturation p	650	32	39.5	400	2	T05563	hypothetical prote
578	32	39.5	197	2	JC2320	hypothetical 22K p	651	32	39.5	401	2	AD1264	ammonium transport
579	32	39.5	209	2	S27494	nods protein - Bra	652	32	39.5	403	1	GRECY	tyrosine-specific
580	32	39.5	214	2	B40296	hypothetical prote	653	32	39.5	403	2	G90955	tyrosine-specific
581	32	39.5	222	2	S63320	probable membrane	654	32	39.5	403	2	D85804	probable membrane
582	32	39.5	222	2	S56195	ATP-binding protei	655	32	39.5	405	2	A75459	sodium extrusion p
583	32	39.5	222	2	AC2397	hypothetical prote	656	32	39.5	412	2	F91068	hypothetical prote
584	32	39.5	225	2	T26491	hypothetical prote	657	32	39.5	412	2	H85912	hypothetical prote
585	32	39.5	226	2	T08335	hypothetical prote	658	32	39.5	413	2	B50162	DNA/pantothenate m
586	32	39.5	233	2	T47136	hypothetical prote	659	32	39.5	413	2	T28646	Y4J5 protein - Rhi
587	32	39.5	238	2	H83371	probable amino aci	660	32	39.5	414	2	T28646	two-component resp
588	32	39.5	238	2	AD2034	hypothetical prote	661	32	39.5	414	2	AI2295	hypothetical prote
589	32	39.5	245	2	C95314	TRm23b IS ATP-bind	662	32	39.5	415	2	C81324	ubiquinol-cytochro
590	32	39.5	245	2	D95334	TRm23b IS ATP-bind	663	32	39.5	416	2	A83133	probable MFS trans
591	32	39.5	256	1	S10164	hydrolase, alpha/b	664	32	39.5	419	2	T24820	hypothetical prote
592	32	39.5	261	2	C87591	indole-3-glycerol-	665	32	39.5	419	2	T24820	hypothetical prote
593	32	39.5	262	1	S17705	hypothetical prote	666	32	39.5	420	2	D83952	hypothetical prote
594	32	39.5	263	1	S43189	probable endochiti	667	32	39.5	421	2	T27311	hypothetical prote
595	32	39.5	265	2	H84867	CT338 hypothet	668	32	39.5	426	2	H87307	ubiquinol-cytochro
596	32	39.5	265	2	G72126	CT338 hypothet	669	32	39.5	432	2	A70714	hypothetical prote
597	32	39.5	265	2	F86495	hypothetical prote	670	32	39.5	432	2	A70714	hypothetical prote
598	32	39.5	265	2	A86548	hypothetical prote	671	32	39.5	438	2	AD1888	hypothetical prote
599	32	39.5	265	2	H72075	hypothetical prote	672	32	39.5	438	2	T51395	hypothetical prote
600	32	39.5	265	2	C81592	hypothetical prote	673	32	39.5	439	2	T01270	probable cadaverin
601	32	39.5	266	2	T16106	hypothetical prote	674	32	39.5	445	2	D82342	glucarate dehydrat
602	32	39.5	268	2	I53029	enkephalin A precu	675	32	39.5	446	2	H65060	glucarate dehydrat
603	32	39.5	268	2	B82929	unknown protein F1	676	32	39.5	447	2	A80094	probable gluconate
604	32	39.5	272	2	C96742	hypothetical prote	677	32	39.5	453	2	AE0164	hypothetical prote
605	32	39.5	274	2	C83709	hypothetical prote	678	32	39.5	456	2	AE0164	probable exported
606	32	39.5	285	2	E75507	hypothetical prote	679	32	39.5	456	2	T47389	hypothetical prote
607	32	39.5	286	2	T41931	minor capsid prote	680	32	39.5	459	2	T45559	guanine nucleotide
608	32	39.5	290	2	T16540	hypothetical prote	681	32	39.5	460	2	G01210	NADH2 dehydrogenas
609	32	39.5	293	2	AB2295	succinyl-CoA synth	682	32	39.5	463	2	A44808	cellulase (EC 3.2.
610	32	39.5	295	2	AE1019	lysr family regula	683	32	39.5	465	2	G50961	flagellar capping
611	32	39.5	297	2	AB1962	acetylglutamate ki	684	32	39.5	465	2	G85809	flagellar capping
612	32	39.5	303	2	S73051	tropinesterase hom	685	32	39.5	468	2	A64956	flagellar hook-ase
613	32	39.5	304	2	H86332	T20H2.26 protein -	686	32	39.5	469	2	AE0835	probable type I se

687	32	39.5	469	2	A87467	glycosyl hydrolase	760	32	39.5	901	1	WMNVTN	104K glycoprotein
688	32	39.5	471	2	AG0146	probable pyridoxal	761	32	39.5	914	2	S18942	hypothetical prote
689	32	39.5	471	2	T32788	hypothetical prote	762	32	39.5	974	2	A71466	probable zinc meta
690	32	39.5	478	2	F82175	conserved hypothet	763	32	39.5	975	2	C81728	metalloproteinase,
691	32	39.5	484	2	AE0887	probable membrane	764	32	39.5	986	2	I78844	receptor protein-c
692	32	39.5	490	2	F87443	conserved hypothet	765	32	39.5	1059	2	E87058	isoletcyl-tRNA syn
693	32	39.5	492	2	T28025	hypothetical prote	766	32	39.5	1071	2	A84785	hypothetical prote
694	32	39.5	494	2	AG3103	mannotol 2-dehydro	767	32	39.5	1089	1	S33727	platelet-derived g
695	32	39.5	494	2	A82294	probable carbon st	768	32	39.5	1098	2	G70697	probable arabinosy
696	32	39.5	496	2	A54757	protein disulfide-	769	32	39.5	1123	2	C96622	probable Afpase F2
697	32	39.5	500	2	T36090	probable integral	770	32	39.5	1139	2	T23018	hypothetical prote
698	32	39.5	500	2	A86112	probable transport	771	32	39.5	1141	2	T05068	hypothetical prote
699	32	39.5	503	2	E83490	probable MFS trans	772	32	39.5	1200	2	F96711	hypothetical prote
700	32	39.5	505	2	E87021	probable integral-	773	32	39.5	1203	2	T04294	hypothetical prote
701	32	39.5	506	2	H70528	hypothetical prote	774	32	39.5	1213	2	D86384	unknown protein [i
702	32	39.5	509	2	S17597	lg delta chain (WI	775	32	39.5	1214	2	T18549	DNA polymerase hom
703	32	39.5	509	2	A82840	beta-lactamase ind	776	32	39.5	1214	2	T30941	DNA polymerase - c
704	32	39.5	509	2	H83002	drug efflux transp	777	32	39.5	1239	2	T06143	disease resistance
705	32	39.5	510	2	H82624	di-tripeptide ABC	778	32	39.5	1244	2	C96584	hypothetical prote
706	32	39.5	511	2	C69199	phenylalanine-tRNA	779	32	39.5	1368	2	S52781	neurocan - mouse
707	32	39.5	514	2	S56384	hypothetical 56.3K	780	32	39.5	1315	2	G95722	hypothetical prote
708	32	39.5	514	2	A91271	probable transport	781	32	39.5	1340	2	AH2219	hypothetical prote
709	32	39.5	516	2	H70935	hypothetical prote	782	32	39.5	1633	2	T01879	hypothetical prote
710	32	39.5	518	2	AD1930	ammonium transport	783	32	39.5	1712	1	CGHU2B	collagen alpha 2(I
711	32	39.5	520	2	T21462	hypothetical prote	784	32	39.5	1793	2	T47897	guanine nucleotide
712	32	39.5	522	2	S77073	hypothetical prote	785	32	39.5	2052	2	T18290	FYVE finger-contai
713	32	39.5	525	2	C98183	mannotol 2-dehydro	786	32	39.5	2149	2	C96695	ribulose biphosph
714	32	39.5	539	2	A35052	interleukin-2 rece	787	32	39.5	2210	1	RRAPLC	genome polyprotein
715	32	39.5	544	2	F81059	YhbX/YhjW/YijP/Yj	788	32	39.5	2241	2	S09811	hypothetical prote
716	32	39.5	547	2	T39478	zinc-finger protei	789	32	39.5	2252	2	S06188	genome polyprotein
717	32	39.5	557	2	F69481	probable acid-CoA	790	32	39.5	2395	1	S50820	surface protein ty
718	32	39.5	564	2	AG2823	ABC transporter, m	791	32	39.5	2703	1	A24420	notch protein - fr
719	32	39.5	564	2	F97601	afub (AE006182) [i	792	32	39.5	2948	2	T22664	hypothetical prote
720	32	39.5	583	2	T20291	hypothetical prote	793	32	39.5	3083	2	AH2493	hypothetical prote
721	32	39.5	598	2	E83593	probable acyl-CoA	794	32	39.5	3149	1	Q0BE8	BPLF1 protein - hu
722	32	39.5	598	2	T51033	hypothetical prote	795	32	39.5	3649	1	S18268	delta-[L]-alpha-ami
723	32	39.5	602	2	A44409	gamma-aminobutyric	796	32	39.5	4717	2	T41581	hypothetical coile
724	32	39.5	602	2	A45078	gamma-aminobutyric	797	31.5	38.9	32	2	S21547	T-cell receptor al
725	32	39.5	615	2	AI0615	probable exported	798	31.5	38.9	105	2	PHI020	Ig heavy chain V r
726	32	39.5	624	2	S50650	hypothetical prote	799	31.5	38.9	113	2	S26263	T-cell receptor be
727	32	39.5	640	2	T28631	Y4CD protein - Rhi	800	31.5	38.9	113	2	S26262	T-cell receptor be
728	32	39.5	650	2	JC7095	sodium-dependent v	801	31.5	38.9	113	2	S17385	T-cell receptor be
729	32	39.5	653	2	A84675	hypothetical prote	802	31.5	38.9	113	2	S26266	T-cell receptor be
730	32	39.5	657	2	C84675	hypothetical prote	803	31.5	38.9	113	2	S17386	T-cell receptor be
731	32	39.5	662	2	T17211	hypothetical prote	804	31.5	38.9	122	2	A83756	hypothetical prote
732	32	39.5	667	2	T09013	RING finger protei	805	31.5	38.9	146	2	S26408	T-cell receptor be
733	32	39.5	698	2	H75577	probable drug tran	806	31.5	38.9	148	2	D87464	conserved hypothet
734	32	39.5	703	2	D84680	ribonucleoside-dip	807	31.5	38.9	152	2	S21826	T-cell receptor be
735	32	39.5	705	2	AG0306	probable ABC trans	808	31.5	38.9	175	2	H38164	hypothetical prote
736	32	39.5	707	2	T23361	hypothetical prote	809	31.5	38.9	182	2	T49852	hypothetical prote
737	32	39.5	708	1	A32959	tryptophan synphas	810	31.5	38.9	274	2	T04619	hypothetical prote
738	32	39.5	720	2	E86297	F309.6 protein - A	811	31.5	38.9	276	2	S74967	UDP-3-0-acyl N-ace
739	32	39.5	742	2	S27263	synaptic vesicle p	812	31.5	38.9	278	1	Q0BE17	dUTP diphosphatase
740	32	39.5	755	1	S74695	hypothetical prote	813	31.5	38.9	278	2	AB3307	glycine betaine/l-
741	32	39.5	755	2	T46411	hypothetical prote	814	31.5	38.9	352	2	T22384	hypothetical prote
742	32	39.5	757	2	B84790	hypothetical prote	815	31.5	38.9	380	1	WZ2MP	dihydrodipicolinat
743	32	39.5	771	2	S72526	inorganic diphosph	816	31.5	38.9	431	2	G71524	probable aspartoki
744	32	39.5	776	1	RGEACR	aerobic respiratio	817	31.5	38.9	437	2	D81681	aspartokinase III
745	32	39.5	778	2	AG0906	aerobic respiratio	818	31.5	38.9	440	2	A55415	rhannogalacturonas
746	32	39.5	778	2	A91140	aerobic respiratio	819	31.5	38.9	466	2	S53713	nitric-oxide reduc
747	32	39.5	778	2	D85985	aerobic respiratio	820	31.5	38.9	505	2	T03394	probable betaine-a
748	32	39.5	781	2	A85035	hypothetical prote	821	31.5	38.9	505	2	A82261	apolipoprotein N-a
749	32	39.5	781	2	T51433	probable cation tr	822	31.5	38.9	522	2	T50359	probable aldehyde-
750	32	39.5	784	2	AH2560	hypothetical prote	823	31.5	38.9	537	2	T02982	probable sucrose t
751	32	39.5	786	2	D75630	glycerophosphoryl	824	31.5	38.9	584	1	VCM5IA	env polyprotein pr
752	32	39.5	788	2	AF0122	probable membrane	825	31.5	38.9	612	2	T13616	hypothetical prote
753	32	39.5	792	2	G84830	probable potassium	826	31.5	38.9	622	2	G81981	probable lipopolys
754	32	39.5	796	2	A45695	capsid precursor -	827	31.5	38.9	623	2	B81037	lipopolysaccharide
755	32	39.5	797	2	E84642	hypothetical prote	828	31.5	38.9	663	2	S69566	glutamine transpor
756	32	39.5	806	2	S76650	sensory transducti	829	31.5	38.9	965	2	T17395	reverse transcript
757	32	39.5	846	2	AI2336	primosomal protein	830	31.5	38.9	1258	2	T14855	probable DEAH Afp-
758	32	39.5	869	2	AD2418	SWI/SNF family hel	831	31	38.3	27	1	MEHB2	melittin, minor -
759	32	39.5	896	2	T07408	lipoxigenase (BC 1	832	31	38.3	39	2	E82725	hypothetical prote

833	31	38.3	46	2	T37121	hypothetical prote	906	31	38.3	229	2	H70660	probable serine ac
834	31	38.3	59	2	C91211	hypothetical prote	907	31	38.3	229	2	C87047	conserved hypothet
835	31	38.3	63	2	T08130	oleosin-like prote	908	31	38.3	229	2	JC7219	nuclear protein SR
836	31	38.3	86	2	G95993	hypothetical prote	909	31	38.3	231	2	B76456	hypothetical prote
837	31	38.3	89	2	C42525	hypothetical prote	910	31	38.3	234	2	S72254	glycerol uptake fa
838	31	38.3	89	2	JQ1822	A-ORF-P protein -	911	31	38.3	237	2	F70606	probable pknM prot
839	31	38.3	89	2	G84976	factor-for-inversi	912	31	38.3	246	2	AF0429	conserved hypothet
840	31	38.3	98	2	D75488	conserved hypothet	913	31	38.3	247	2	T10524	tonoplast intrinsi
841	31	38.3	99	2	S57386	hypothetical prote	914	31	38.3	249	2	JQ2287	SFC1 protein - so
842	31	38.3	106	4	PL0270	hypothetical prote	915	31	38.3	249	2	T09297	tonoplast intrinsi
843	31	38.3	107	2	PL0272	ig kappa chain v r	916	31	38.3	249	2	H83215	probable ATP-bindi
844	31	38.3	107	2	PL0271	ig kappa chain v r	917	31	38.3	249	2	H71131	hypothetical prote
845	31	38.3	107	2	PL0269	ig kappa chain v r	918	31	38.3	251	2	S22202	tonoplast intrinsi
846	31	38.3	109	2	S74284	hypothetical prote	919	31	38.3	252	2	F83443	hypothetical prote
847	31	38.3	110	2	S13688	ig heavy chain v r	920	31	38.3	252	2	A71691	hypothetical prote
848	31	38.3	111	2	S13687	ig heavy chain v r	921	31	38.3	253	2	S13718	probable membrane
849	31	38.3	112	2	S13686	ig heavy chain v r	922	31	38.3	253	2	S15189	gamma tonoplast in
850	31	38.3	112	2	S13685	ig heavy chain v r	923	31	38.3	253	2	D87392	nucleotidyltransfe
851	31	38.3	112	2	S51123	genome polypotein	924	31	38.3	254	2	AB0655	phosphatidylglycer
852	31	38.3	122	4	T01781	probable pol prote	925	31	38.3	254	2	B72549	hypothetical prote
853	31	38.3	123	2	G82357	conserved hypothet	926	31	38.3	255	2	JQ2288	SFC2 protein - so
854	31	38.3	124	2	B27664	T-cell receptor be	927	31	38.3	255	2	T04389	NBS-LRR type resis
855	31	38.3	124	2	T32823	hypothetical prote	928	31	38.3	255	2	H97757	hydrolase homolog
856	31	38.3	127	2	C75315	diacylglycerol kin	929	31	38.3	256	2	A13589	pca regulon regula
857	31	38.3	127	2	AF1895	hypothetical prote	930	31	38.3	256	2	D86544	NADH (ubiquinone)
858	31	38.3	130	2	A26316	hypothetical prote	931	31	38.3	256	2	A96027	probable ABC trans
859	31	38.3	133	2	F81783	hypothetical prote	932	31	38.3	256	2	B83833	hypothetical prote
860	31	38.3	133	2	D81174	hypothetical prote	933	31	38.3	260	2	C86702	proliporotein dia
861	31	38.3	134	1	WTBO	seminal fluid prot	934	31	38.3	261	2	JCS806	aquaporin 8 - mous
862	31	38.3	140	2	S54245	ig mu heavy chain	935	31	38.3	261	2	A42201	bacteriorhodopsin
863	31	38.3	140	2	G83187	conserved hypothet	936	31	38.3	262	2	I40221	diviB protein - Ba
864	31	38.3	141	2	G27557	T-cell receptor be	937	31	38.3	262	2	B70904	hypothetical prote
865	31	38.3	142	2	A27608	T-cell receptor be	938	31	38.3	265	2	AH0755	conserved hypothet
866	31	38.3	143	2	E96638	hypothetical prote	939	31	38.3	265	2	D82581	probable signal pe
867	31	38.3	146	2	E70557	hypothetical prote	940	31	38.3	266	2	A29310	MHC class II hista
868	31	38.3	147	2	T04470	probable trypsin i	941	31	38.3	266	2	H75429	transcription regu
869	31	38.3	153	2	D24372	dnpD protein - Str	942	31	38.3	267	2	B32410	mastocytoma protei
870	31	38.3	153	2	F98079	restriction syatem	943	31	38.3	269	2	B32410	pyrroline-5-carbox
871	31	38.3	153	2	H95215	DnpD protein [mpo	944	31	38.3	270	2	G97299	indole-3-glycerol-
872	31	38.3	160	2	H72569	hypothetical prote	945	31	38.3	270	2	T46856	hypothetical prote
873	31	38.3	163	1	E69231	conserved hypothet	946	31	38.3	272	2	T35231	hypothetical prote
874	31	38.3	167	2	T45261	hypothetical prote	947	31	38.3	273	2	T11706	NADH2 dehydrogenas
875	31	38.3	169	2	B26395	T-cell receptor be	948	31	38.3	274	2	T11727	NADH2 dehydrogenas
876	31	38.3	173	2	H87629	conserved hypothet	949	31	38.3	274	2	T25783	hypothetical prote
877	31	38.3	173	2	G96005	probable RNA polym	950	31	38.3	275	2	D83841	NADH2 dehydrogenas
878	31	38.3	174	2	A49181	alpha B-crystallin	951	31	38.3	276	2	D56406	purine nucleoside
879	31	38.3	174	2	S74666	DnaJ protein - Syn	952	31	38.3	277	2	S57381	casein kinase I (E
880	31	38.3	176	2	S45343	glomerulosclerosis	953	31	38.3	278	2	B83354	protein disulfide
881	31	38.3	176	2	S29031	mpv17 protein - mo	954	31	38.3	278	2	A71498	hypothetical prote
882	31	38.3	177	2	H87412	hypothetical prote	955	31	38.3	279	2	T25801	hypothetical prote
883	31	38.3	178	2	JQ1547	stripe disease-spe	956	31	38.3	280	2	E87429	conserved hypothet
884	31	38.3	181	2	A37878	phosphoribosylglyc	957	31	38.3	284	2	S27931	Env/v-mpl fusion p
885	31	38.3	181	2	D95005	phosphoribosylglyc	958	31	38.3	286	2	B83343	hypothetical prote
886	31	38.3	182	2	C84982	hypothetical prote	959	31	38.3	286	2	A54026	hypothetical prote
887	31	38.3	182	2	D82240	conserved hypothet	960	31	38.3	289	2	F69313	alcohol sulfotrans
888	31	38.3	188	2	D82240	hypothetical prote	961	31	38.3	289	2	C83207	probable hydrolase
889	31	38.3	190	2	F87263	alkyl hydroperoxid	962	31	38.3	289	2	H81729	conserved hypothet
890	31	38.3	195	2	F69283	transcription regu	963	31	38.3	291	2	H75055	succinyl-CoA synth
891	31	38.3	197	2	G32973	probable transcrip	964	31	38.3	292	2	T43930	hypothetical prote
892	31	38.3	202	2	T36918	probable membrane	965	31	38.3	295	2	S01441	class II histocomp
893	31	38.3	202	2	S56267	hypothetical prote	966	31	38.3	296	2	AH0341	probable aldo/keto
894	31	38.3	203	2	F51098	hypothetical prote	967	31	38.3	297	2	G70887	fibpC1 protein prec
895	31	38.3	209	2	S51098	probable alkyl hyd	968	31	38.3	299	2	AG0603	hypothetical ABC t
896	31	38.3	209	2	JX0266	platelet aggregat	969	31	38.3	303	2	H64820	probable oligopept
897	31	38.3	211	2	C81001	probable pyrazinam	970	31	38.3	303	2	G90742	probable transport
898	31	38.3	211	2	H82018	probable nicotinam	971	31	38.3	303	2	B85593	probable transport
899	31	38.3	212	2	B83283	probable transcrip	972	31	38.3	304	2	B69474	endoglucanase (cel
900	31	38.3	212	2	S37973	DNA-directed RNA p	973	31	38.3	304	2	A87258	hypothetical prote
901	31	38.3	214	2	S59149	H+-transporting tw	974	31	38.3	307	2	AB3614	glycosyl transfera
902	31	38.3	216	2	JX0265	platelet aggregat	975	31	38.3	307	2	T25875	hypothetical prote
903	31	38.3	217	2	H95995	probable isochoris	976	31	38.3	309	2	G81381	probable integral
904	31	38.3	218	2	JC7220	nuclear protein SR	977	31	38.3	310	2	S17177	probable G protein
905	31	38.3	224	2	T40315	hypothetical prote	978	31	38.3	320	2	A46152	A3 adenosine recep
			226	2	B69079	alkyl hydroperoxid							

979 31 38.3 323 2 E71134 hypothetical prote
980 31 38.3 324 2 T07832 probable steroid s
981 31 38.3 328 2 S64098 nucleotide-binding
982 31 38.3 328 2 T37734 SUR4 family protei
983 31 38.3 329 2 S74595 hypothetical prote
984 31 38.3 331 2 B48445 glyceraldehide-3-p
985 31 38.3 331 2 A71870 integrase/recombin
986 31 38.3 331 2 J70589 pectinesterase (EC
987 31 38.3 331 2 T25785 hypothetical prote
988 31 38.3 332 2 F75473 conserved hypothet
989 31 38.3 333 2 T65989 G protein-coupled
990 31 38.3 333 2 S12392 fanF protein - Esc
991 31 38.3 333 2 F90172 hypothetical prote
992 31 38.3 334 2 AD0775 probable hydrolase
993 31 38.3 335 2 T25498 hypothetical prote
994 31 38.3 337 2 S68678 adenosine receptor
995 31 38.3 340 2 T13781 NADH2 dehydrogenas
996 31 38.3 342 2 T29245 hypothetical prote
997 31 38.3 343 2 AG2465 hypothetical prote
998 31 38.3 344 2 D97899 hypothetical prote
999 31 38.3 345 1 JH0185 D-amino-acid oxida
1000 31 38.3 345 2 B69678 involved in polyke

ALIGNMENTS

RESULT 1
T30343
irp3 protein - Yersinia enterocolitica
C;Species: Yersinia enterocolitica
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30343
J. Pelludat, C.; Rakin, A.; Jacobi, S.; Schubert, S.; Heesemann, J.
R. Bacteriol. 180, 538-546, 1998
A;Title: The Yersiniaabactin biosynthetic gene cluster of Yersinia enterocolitica: organi
A;Reference number: 220833; MUID:98117033; PMID:19457855
A;Accession: T30343
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-365 <PBL>
A;Cross-references: UNIPROT:O54512; UNIPARC:UPI00000B099E; EMBL:Y12527; PIDN:CAA73128.1
C;Genetics:
A;Gene: irp3

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWL 12
||| ||| :|||
Db 290 CETVGPBGVSWL 301

RESULT 2
T17441
hypothetical protein ybtU [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17441; T47049
R;Petherston, J.D.; Bertolino, V.J.; Perry, R.D.
Mol. Microbiol. 32, 289-299, 1999
A;Title: YbtP and YbtQ: two ABC transporters required for iron uptake in Yersinia pestis
A;Reference number: Z18782; MUID:99248409; PMID:10231486
A;Accession: T17441
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-365 <PBL>
A;Cross-references: UNIPROT:Q923C6; UNIPARC:UPI00000B43BB; EMBL:AF091251; NID:g3818595;
R;Buchrieser, C.; Ruenikok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunat, F.; Carniel
submitted to the EMBL Data Library, October 1998
A;Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A;Reference number: Z24348

A;Accession: T47049
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-365 <BUC>
A;Cross-references: UNIPARC:UPI00000B43BB; EMBL:AL031866; PIDN:CAA21392.1
A;Experimental source: strain 6/69
C;Genetics:
A;Gene: ybtU

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWL 12
||| ||| :|||
Db 290 CETVGPBGVSWL 301

RESULT 3
AI0232
Yersiniaabactin biosynthetic protein YbtU [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AI0232
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0232
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-366 <KUR>
A;Cross-references: UNIPROT:Q8ZF25; UNIPARC:UPI00000CD874; GB:AL590842; PIDN:CAC90725.1;
C;Genetics:
A;Gene: ybtU

Query Match 56.8%; Score 46; DB 2; Length 366;
Best Local Similarity 66.7%; Pred. No. 4.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPMSLMSWL 12
||| ||| :|||
Db 291 CETVGPBGVSWL 302

RESULT 4
ABI885
hypothetical protein alr0627 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: ABI885
R;Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.;
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: ABI807; MUID:21595285; PMID:11759840
A;Accession: ABI885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-907 <KUR>
A;Cross-references: UNIPROT:Q8YZ60; UNIPARC:UPI00000CDEDB; GB:BA000019; PIDN:BA072585.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0627

Query Match 55.6%; Score 45; DB 2; Length 907;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VGPMSLMSWL 12


```
Db      106 IGPTLSWL 114
      :||:|||||
RESULT 5
A35098
MHC class III histocompatibility antigen HLA-B-associated transcript 3 - human
C;Species: Homo sapiens (man)
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 31-Dec-2004
C;Accession: A35098
R;Banerji, J.; Sands, J.; Strominger, J.L.; Spies, T.
Proc. Natl. Acad. Sci. U.S.A. 87, 2374-2378, 1990
A;Title: A gene pair from the human major histocompatibility complex encodes large proli
A;Reference number: A35098; MUID:90192810; PMID:2156268
A;Accession: A35098
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1132 <BAN>
A;Cross-references: UNIPROT:P46379; UNIPARC:UPI00001267A0; GB:M33519; NID:g179346; PIDN:
F;17-91/Domain: ubiquitin homology <UB>

      Query Match      54.3%; Score 44; DB 2; Length 1132;
      Best Local Similarity 80.0%; Pred. No. 31;
      Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VGPSLSWLT 13
      | |||:|||||
Db      906 VNPSLSWLT 915

RESULT 6
T35918
hypothetical protein SC982.22c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T35918
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z21593
A;Accession: T35918
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-224 <SAU>
A;Cross-references: UNIPROT:Q9Z537; UNIPARC:UPI00000DAEBB; EMBL:AL035212; PIDN:CAA22802.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC982B:SC982.22c

      Query Match      53.1%; Score 43; DB 2; Length 224;
      Best Local Similarity 72.7%; Pred. No. 9.8;
      Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 ELVGPSLSWL 12
      |||||:|
Db      162 ELVGPELWPL 172

RESULT 7
C55581
naad protein - Klebsiella pneumoniae
C;Species: Klebsiella pneumoniae
C;Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 05-Oct-2004
C;Accession: C55581
R;Lin, J.T.; Goldman, B.S.; Stewart, V.
J. Bacteriol. 176, 2551-2559, 1994
A;Title: The nasFEDCEBA operon for nitrate and nitrite assimilation in Klebsiella pneumon
A;Reference number: A55581; MUID:94222832; PMID:8169203
A;Accession: C55581
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-261 <LIN>
A;Cross-references: UNIPARC:UPI000017A0F2; GB:L27431; NID:g1119218; PID:g473439
C;Keywords: ATP; nucleotide binding; P-loop
```

F;23-214/Domain: ATP-binding cassette homology <ABC>
F;40-47/Region: nucleotide-binding motif A (P-loop)

Query Match 52.5%; Score 42.5; DB 2; Length 261;
Best Local Similarity 37.5%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 2; Indels 11; Gaps 1;

Qy 2 ELVGP-----SLMSWLT 14

Db 70 EIAGPGPERAVVFQNHSLLPWLT 93

RESULT 8

T00424

probable Na⁺/Ca²⁺ antiporter [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein T30B22.10

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00424; C84917

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Maso

submitted to the EMBL Data Library, October 1998

A;Description: Arabidopsis thaliana chromosome II BAC T30B22 genomic sequence.

A;Reference number: Z14149

A;Accession: T00424

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-538 <ROU>

A;Cross-references: UNIPROT:O2252; UNIPARC:UPI000017AFBD; EMBL:AC002535; NID:g2529657;

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, J.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84917

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-538 <STO>

A;Cross-references: UNIPARC:UPI000017AFBD; GB:AB002093; NID:g3522931; PIDN:AAC62871.1;

C;Genetics:

A;Gene: At2g47600; T30B22.10

A;Map position: 2

A;Introns: 131/1; 189/3; 229/2; 287/1; 333/1; 397/1; 440/2

Query Match 51.9%; Score 42; DB 2; Length 538;

Best Local Similarity 66.7%; Pred. No. 33;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELVGPSLSWLT 13

Db 504 ELGGPELWALT 515

RESULT 9

S33475

hypothetical protein - Trypanosoma brucei

C;Species: Trypanosoma brucei

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S33475

R;Gottesdiener, K.M.

submitted to the EMBL Data Library, May 1993

A;Description: A new USG expression site associated gene in the promoter region.

A;Reference number: S33475

A;Accession: S33475

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-686 <GOT>

A;Cross-references: UNIPROT:Q07492; UNIPARC:UPI0000077EA1; EMBL:X73565; NID:g311493; PI

Query Match 51.9%; Score 42; DB 2; Length 686;

Best Local Similarity 70.0%; Pred. No. 42;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LVGPSLSWSL 12
Db 257 LAGPSLVSWV 266

RESULT 10
G88065
protein T16A1.2 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: G88065
R/anonymous: The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: G88065
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-786 <STO>
A/Cross-references: UNIPROT:P91451; UNIPARC:UPI000007BD2C; GB:chr_II; PIDN:AB37879.1; PI
C/Genetics:
A/Gene: T16A1.2
A/Map position: 2

Query Match 51.9%; Score 42; DB 2; Length 786;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 GPSLSWSLTC 14
Db 466 GKSLLSWLKC 475

RESULT 11
S49844
Probable membrane protein YDR089w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein D4495; hypothetical protein YD6652.01; hypothetical
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1995 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C/Accession: S49844; S48778; S55838; S67906; S58089
R/Richards, C.; Harris, D.E.
Submitted to the EMBL Data Library, November 1994
A/Reference number: S49823
A/Accession: S49844
A/Molecule type: DNA
A/Residues: 1-260 <RIC>
A/Cross-references: UNIPROT:P38966; UNIPARC:UPI0000168A94; EMBL:Z46796; NID:G577794; PI
R/Coster, F.; Jonniaux, J.L.; Goffeau, A.
Submitted to the EMBL Data Library, October 1994
A/Reference number: S48778
A/Accession: S48778
A/Molecule type: DNA
A/Residues: 1-161 <COS>
A/Cross-references: UNIPARC:UPI0000168B72; EMBL:X82086; NID:G558241; PID:G558262
R/Coster, F.; Jonniaux, J.L.; Goffeau, A.
Yeast 11, 673-679, 1995
A/Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr
A/Reference number: S55819; MUID:56093910; PMID:7483840
A/Accession: S55838
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-161 <COS>
A/Cross-references: UNIPARC:UPI0000168B72; EMBL:X82086; NID:G558241; PIDN:CAA57618.1; PI
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
R/Foury, F.; Jonniaux, J.L.; Purnelle, B.; Coster, F.; Goffeau, A.
Submitted to the Protein Sequence Database, July 1996
A/Reference number: S67889
A/Accession: S67906
A/Molecule type: DNA
A/Residues: 1-161 <FOU>

A/Cross-references: UNIPARC:UPI0000168B72; EMBL:Z74385; MIPS:YDR089w
A/Experimental source: strain S288C
R/Oliver, K.; Harris, D.
Submitted to the EMBL Data Library, July 1995
A/Reference number: S58089
A/Accession: S58089
A/Molecule type: DNA
A/Residues: 249-869 <OLI>
A/Cross-references: UNIPARC:UPI000006B647; EMBL:Z50111; NID:G914872; PID:G914873
A/Experimental source: strain AB972
C/Genetics:
A/Cross-references: SGD:S0002496
A/Map position: 4R
C/Superfamily: Saccharomyces cerevisiae probable membrane protein YDR089w
C/Keywords: transmembrane protein
F/777-793/Domain: transmembrane #status predicted <TM1>
F/811-827/Domain: transmembrane #status predicted <TM2>
F/828-844/Domain: transmembrane #status predicted <TM3>

Query Match 51.9%; Score 42; DB 2; Length 869;
Best Local Similarity 42.9%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CELVGPSLSWSLTC 14
Db 537 CSLPPDLTTWKIC 550

RESULT 12
EQBOA
enkephalin e-12 precursor - bovine
N/Alternate names: proenkephalin A precursor
N/Contains: Leu-enkephalin; Met-enkephalin; Met-enkephalin-Arg-Gly-Leu; Met-enkephalin-A
C/Species: Bos primigenius taurus (cattle)
C/Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C/Accession: A93272; A93273; A33304; A19448; S62377; S13473; S13474; I45950; A01475
R/Noda, M.; Furutani, Y.; Takahashi, H.; Toyosato, M.; Hirose, T.; Inayama, S.; Nakanishi
Nature 295, 202-206, 1982
A/Title: Cloning and sequence analysis of cDNA for bovine adrenal preproenkephalin.
A/Reference number: A93272; MUID:82125409; PMID:6276759
A/Accession: A93272
A/Molecule type: mRNA
A/Residues: 1-263 <NOD>
A/Cross-references: UNIPROT:P01211; UNIPARC:UPI00001315A6; GB:J00005; GB:V00109; NID:G16:
A/Experimental source: adrenal medulla
R/Gubler, U.; Seeburg, P.; Hoffman, B.J.; Gage, L.P.; Udenfriend, S.
Nature 295, 206-208, 1982
A/Title: Molecular cloning establishes proenkephalin as precursor of enkephalin-contains
A/Reference number: A93273; MUID:82125410; PMID:6173760
A/Accession: A93273
A/Molecule type: mRNA
A/Residues: 25-263 <GUB>
A/Cross-references: UNIPARC:UPI0000173414
A/Experimental source: adrenal medulla
R/Watkinson, A.; Young, J.; Varro, A.; Dockray, G.J.
J. Biol. Chem. 264, 3061-3065, 1989
A/Title: The isolation and chemical characterization of phosphorylated enkephalin-contains
A/Reference number: A33304; MUID:89123418; PMID:2914943
A/Accession: A33304
A/Molecule type: protein
A/Residues: 233-251 <WAT>
A/Cross-references: UNIPARC:UPI0000173415
A/Experimental source: adrenal medulla
R/Jones, B.N.; Shively, J.E.; Kilpatrick, D.L.; Stern, A.S.; Lewis, R.V.; Kojima, K.; Ude
Proc. Natl. Acad. Sci. U.S.A. 79, 2096-2100, 1982
A/Title: Adrenal opioid proteins of 8600 and 12,600 daltons: intermediates in proenkeph
A/Reference number: A19448; MUID:82197644; PMID:6952256
A/Accession: A19448
A/Molecule type: protein
A/Residues: 25-60;61-62, 'Z', 64-66; 'Z', 68-69, 'Z', 71-73;74, 'Z', 76-78, 'B', 80-86; 'ZZZ', 90-94,
A/Cross-references: UNIPARC:UPI0000173416; UNIPARC:UPI0000173417; UNIPARC:UPI0000173418;
R/Goumon, Y.; Strub, J.M.; Moniatte, M.; Nullans, G.; Poteur, L.; Hubert, P.; van Dorssel
Eur. J. Biochem. 235, 516-525, 1996

A;Title: The C-terminal bisphosphorylated proenkephalin-A- (209-237)-peptide from adrenal
A;Reference number: S62377; MUID:96184524; PMID:8654396
A;Accession: S62377
A;Molecule type: protein
A;Residues: 233-261 <GOU>
A;Cross-references: UNIPARC:UPI000017341C
R;Windroiu, T.; Carretero, O.A.; Walz, D.; Scicli, A.G.
Biochim. Biophys. Acta 1076, 9-14, 1991
A;Title: Tissue kallikrein processes small proenkephalin peptides.
A;Reference number: S13473; MUID:91098273; PMID:1986798
A;Accession: S13473
A;Molecule type: protein
A;Residues: 104-137 <MIN>
A;Cross-references: UNIPARC:UPI00000352EE
A;Accession: S13474
A;Molecule type: protein
A;Residues: 206-227 <MIW>
A;Cross-references: UNIPARC:UPI00000352EE
R;Comb, M.; Herbert, E.; Crea, R.
Proc. Natl. Acad. Sci. U.S.A. 79, 360-364, 1982
A;Title: partial characterization of the mrna that codes for enkephalins in bovine adren
A;Reference number: 145950; MUID:82197495; PMID:6952189
A;Accession: 145950
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 211-229 <COM>
A;Cross-references: UNIPARC:UPI000016C34D; GB:J00012; NID:G163435; PIDN:AAA30673.1; PID:
C;Superfamily: proenkephalin
C;Keywords: neuropeptide; opioid peptide; phosphoprotein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;97-101/Product: Met-enkephalin #status experimental <ME1>
F;104-108/Product: Met-enkephalin #status experimental <ME2>
F;133-137/Product: Met-enkephalin #status experimental <ME3>
F;182-189/Product: Met-enkephalin-Arg-Gly-Leu #status predicted <ME1>
F;206-210/Product: Met-enkephalin #status predicted <ME4>
F;226-230/Product: Leu-enkephalin #status predicted <LE1>
F;257-263/Product: Met-enkephalin-Arg-Phe #status predicted <MEF>
F;239,245,247/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 50.6%; Score 41; DB 1; Length 263;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CELVGPMSLWTC 14
||| ||| ||| |||
Db 52 CEGKLPSLKTWETC 65

RESULT 13
A87380
conserved hypothetical protein CC1053 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C;Accession: A87380
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87380
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <STO>
A;Cross-references: UNIPROT:Q9A9D5; UNIPARC:UPI00000C725F; GB:AE005673; NID:G13422351; F
C;Genetics:
A;Gene: CC1053
C;Superfamily: Alpha/beta hydrolase

Query Match 50.6%; Score 41; DB 2; Length 290;
Best Local Similarity 58.3%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPMSLWTC 14
||| ||| ||| |||
Db 98 LVGSMGGWLS 109

RESULT 14
H71631
queine tRNA-ribosyltransferase (tgt) RP721 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: H71631
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: H71631
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <AND>
A;Cross-references: UNIPROT:Q9ZCK8; UNIPARC:UPI0000136CF9; GB:AJ235273; GB:AJ235269; NID
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: tgt; RP721
C;Superfamily: queine tRNA-ribosyltransferase

Query Match 50.6%; Score 41; DB 2; Length 361;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ELVGPMSLW 11
||| ||| ||| |||
Db 325 EILGPMMLTW 334

RESULT 15
S26414
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004
C;Accession: S26414
R;Buffard, D.; Breda, C.; El Turk, J.; Sallaud, C.; Kondorosi, A.; Esnault, R.
submitted to the EMBL Data Library, August 1992
A;Description: Molecular cloning of two chalcone synthase cDNA from alfalfa.
A;Reference number: S26414
A;Accession: S26414
A;Molecule type: mRNA
A;Residues: 1-389 <BUF>
A;Cross-references: UNIPROT:P51078; UNIPARC:UPI0000127922; EMBL:X68106; NID:G19590; PIDN
C;Superfamily: type III polyketide synthase
C;Keywords: acyltransferase; coenzyme A

Query Match 50.6%; Score 41; DB 2; Length 389;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CELVGPMSLW 11
||| ||| ||| |||
Db 84 CEIMAPSLDAW 94

RESULT 16
T03888
Na+/Ca2+,K+-exchanging protein homolog C13D9.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03888
R;Nelson, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid C13D9.
A;Reference number: Z15128
A;Accession: T03888
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-703 <NEL>
A;Cross-references: UNIPROT:O16241; UNIPARC:UPI000007B85B; EMBL:AF016420; NID:g2291168;
C;Genetics:
A;Map position: V
A;Note: C13D9.7

Query Match 50.6%; Score 41; DB 2; Length 703;
Best Local Similarity 53.8%; Pred. No. 63;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWLTLC 14
Db 579 ELLGLTWSMSDC 591
||| :|||
||| :|||

RESULT 17
T34414
hypothetical protein F07E5.8 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34414
R;Du, Z.; Goela, D.
Submitted to the EMBL Data Library, December 1996
A;Description: The sequence of *C. elegans* cosmid F07E5.
A;Reference number: Z21520
A;Accession: T34414
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-760 <DUZ>
A;Cross-references: UNIPROT:P91227; UNIPARC:UPI000007B004; EMBL:U80837; PIDN:AAB37907.1;
A;Experimental source: strain Bristol N2; clone F07E5
C;Genetics:
A;Gene: CESP:F07E5.8
A;Map position: 2
A;Introns: 7/1; 720/2

Query Match 50.6%; Score 41; DB 2; Length 760;
Best Local Similarity 70.0%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GPSLMSWLTLC 14
Db 440 GRSLLSWLXC 449
||| :|||
||| :|||

RESULT 18
AB2301
hypothetical protein asr3961 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AB2301
R;KaneKO, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2301
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84 <KUR>
A;Cross-references: UNIPROT:Q8YQ74; UNIPARC:UPI00000CB97D; GB:BA000019; PIDN:BA075660.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asr3961

Query Match 49.4%; Score 40; DB 2; Length 84;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWLTLC 14
Db ||| :|||
||| :|||

Db 64 CRLGTPPQILWLTC 77

RESULT 19
T02735
hypothetical protein At2g29050 [imported] - *Arabidopsis thaliana*
N;Alternate names: hypothetical protein T914.13
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Oct-2004
C;Accession: T02735; H84691
R;Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Ron
submitted to the EMBL Data Library, August 1998
A;Description: *Arabidopsis thaliana* chromosome II BAC T914 genomic sequence.
A;Reference number: Z14710
A;Accession: T02735
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-372 <ROU>
A;Cross-references: UNIPROT:O81073; UNIPARC:UPI00000AB3DD; EMBL:AC005315; NID:g3461834;
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-372 <GTO>
A;Cross-references: UNIPARC:UPI00000AB3DD; GB:AE002093; NID:g3461845; PIDN:AAC33231.1; G;
C;Genetics:
A;Gene: T914.13; At2g29050
A;Map position: 2
A;Introns: 137/2; 192/1
C;Superfamily: AAAA protein

Query Match 49.4%; Score 40; DB 2; Length 372;
Best Local Similarity 54.5%; Pred. No. 50;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWLTLC 13
Db 104 LLGPSSLTWWT 114
||| :|||
||| :|||

RESULT 20
DEPSXA
3-methyl-2-oxobutanoate dehydrogenase (lipoamide) (EC 1.2.4.4) alpha chain - *Pseudomonas*
N;Alternate names: 2-oxoisovalerate dehydrogenase (lipoamide) E1-alpha chain; branched-c
C;Species: *Pseudomonas putida*
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: S01317; B36133; S63475
R;Burns, G.; Brown, T.; Hatter, K.; Idriss, J.M.; Sokatch, J.R.
Eur. J. Biochem. 176, 311-317, 1988
A;Title: Similarity of the E1 subunits of branched-chain-oxoacid dehydrogenase from *Pseu*
A;Reference number: S01317; MUID:88329084; PMID:3416875
A;Accession: S01317
A;Molecule type: DNA
A;Residues: 1-410 <BUR>
A;Cross-references: UNIPROT:P09060; UNIPARC:UPI0000172024; EMBL:X13004
R;Madhusudhan, K.T.; Huang, G.; Burns, G.; Sokatch, J.R.
J. Bacteriol. 172, 5655-5663, 1990
A;Title: Transcriptional analysis of the promoter region of the *Pseudomonas putida* bran
A;Reference number: A36133; MUID:91008935; PMID:2211503
A;Accession: B36133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-17 <MAD>
A;Cross-references: UNIPARC:UPI0000172025; GB:M33715
R;Hester, K.; Luo, J.; Burns, G.; Braswell, E.H.; Sokatch, J.R.
Eur. J. Biochem. 233, 828-836, 1995
A;Title: Purification of active E1-alpha(2)-beta(2) of *Pseudomonas putida* branched-chain-

A;Reference number: S63475; MUID:96085147; PMID:8521848
A;Accession: S63475
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <HES>
A;Cross-references: UNIPARC:UPI0000172026
C;Genetics:
A;Gene: bkdA1
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bir
C;Keywords: lipoamide; oxidoreductase; phosphoprotein; thiamin pyrophosphate
F;2-410/Product: 3-methyl-2-oxobutanate dehydrogenase (lipoamide) alpha chain #status F
F;202-251/Domain: thiamin pyrophosphate-binding domain homology <TPB>
F;313/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 49.4%; Score 40; DB 1; Length 410;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLMSWLT 13
Db 297 LGPSLIEWVT 306

RESULT 21
C83365
2-Oxoisovalerate dehydrogenase (alpha subunit) PA2247 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: C83365
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: C83365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-410 <STO>
A;Cross-references: UNIPROT:Q91LM2; UNIPARC:UPI00000D74B0; GB:AE004650; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: bkdA1; PA2247
C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-bir

Query Match 49.4%; Score 40; DB 2; Length 410;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLMSWLT 13
Db 297 LGPSLIEWVT 306

RESULT 22
T31134
naphthalene dioxygenase (EC 1.14.12.-) large chain - Spingomonas aromaticivorans plasmid
C;Species: Spingomonas aromaticivorans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
R;Accession: T31134
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A;Description: Complete sequence of a 184 kb catabolic plasmid from Spingomonas aromati
A;Reference number: Z20992
A;Accession: T31134
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-459 <ROM>
A;Cross-references: UNIPROT:O85843; UNIPARC:UPI000005C95F; EMBL:AF079317; NID:G3378261;
C;Genetics:
A;Gene: bphA1
A;Genome: plasmid pNL1

C/Species: Yersinia enterocolitica
C/Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 05-Oct-2004
C/Accession: S54440
R/Stojiljkovic, I.; Hantke, K.
Mol. Microbiol. 13, 719-732, 1994
A/Title: Transport of haemin across the cytoplasmic membrane through a haemin-specific P-Loop
A/Reference number: S54436; MUID:95089707; PMID:7997183
A/Accession: S54440
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-266 <STO>
A/Cross-references: UNIPROT:P74981; UNIPARC:UPI00000B721D; EMBL:X77867; NID:gl619622; P-Loop
A/Note: the authors translated the initiation codon GTG for residue 1 as Val
C/Genetics:
A/Gene: hemV
A/Start codon: GTG
C/Function:
A/Description: component of the periplasmic binding-protein-dependent transport system (C)
C/Keywords: ATP; iron transport; nucleotide binding; P-loop
F/27-224/Domain: ATP-binding cassette homology <ABC>
F/44-51/Region: nucleotide-binding motif A (P-loop)

Query Match 48.1%; Score 39; DB 2; Length 266;
Best Local Similarity 54.5%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CELVGPISLSMW 11
|:|:|:|
Db 68 CQLLGKNSW 78

RESULT 26
C85224
probable LTR retrotransposon (partial) [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: C85224
R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083486; PMID:10617198
A/Accession: C85224
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-306 <STO>
A/Cross-references: UNIPROT:O81859; UNIPARC:UPI000009EF90; GB:NC_001268; NID:g7268774; P-Loop
C/Genetics:
A/Gene: Atg19780
A/Map position: 4
C/Superfamily: retrovirus-related polyprotein

Query Match 48.1%; Score 39; DB 2; Length 306;
Best Local Similarity 54.5%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CELVGPISLSMW 11
|:|:|:|
Db 168 CMFIGDSLVS 178

RESULT 27
T04759
hypothetical protein TIGH5.140 - Arabidopsis thaliana (fragment)
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04759
R/Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A/Reference number: Z15383
A/Accession: T04759
A/Molecule type: DNA
A/Residues: 1-306 <BEV>
A/Cross-references: UNIPROT:O81859; UNIPARC:UPI000009EF90; EMBL:AL024486

C:Function:
A:Description: urea transport
C:Keywords: transmembrane protein; urea transport

Query Match 48.1%; Score 39; DB 2; Length 735;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GPSLMSWLTC 14
| : : : : |
Db 467 GLAIMSWLVC 476

RESULT 30
S75152
hypothetical protein slr1767 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75152
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <KAN>
A:Cross-references: UNIPROT:P73045; UNIPARC:UPI00000C0C7F; EMBL:D90903; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 46.9%; Score 38; DB 2; Length 112;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CELVGPLMSWLTC 14
| : : : : |
Db 64 CRLGIPQIWLTC 77

RESULT 31
A46097
GPI-anchor biosynthesis protein PIG-F - human
C:Species: *Homo sapiens* (man)
C:Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004
C:Accession: A46097
R:Inoue, N.; Kinoshita, T.; Orii, T.; Takeda, J.
J. Biol. Chem. 268, 6882-6885, 1993
A:Title: Cloning of a human gene, PIG-F, a component of glycosyl-phosphatidylinositol an
A:Reference number: A46097; MUID:93216618; PMID:8463218
A:Accession: A46097
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-219 <INO>
A:Cross-references: UNIPROT:Q07326; UNIPARC:UPI0000030AE2; GB:D13435; NID:G303615; PIDN:
C:Genetics:
A:Gene: GDB:PIGF
A:Cross-references: GDB:I38466; OMIM:600153
A:Map position: 2p16-2p21

Query Match 46.9%; Score 38; DB 2; Length 219;
Best Local Similarity 60.0%; Pred. No. 66;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
| : : : : |
Db 130 LLGPNLKAWL 139

RESULT 32
C90848

hypothetical protein EC61755 [imported] - *Escherichia coli* (strain O157:H7, substrain RIN
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C90848
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <HAY>
A:Cross-references: UNIPROT:Q8XC87; UNIPARC:UPI00000D04C2; GB:BA000007; PIDN:BA035178.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC61755
C:Superfamily: yciC protein

Query Match 46.9%; Score 38; DB 2; Length 247;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
| : : : : |
Db 189 LVAPAVLSWL 198

RESULT 33
B85706
hypothetical protein yciC [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85706
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85706
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: UNIPROT:Q8XC87; UNIPARC:UPI00000D04C2; GB:AE005174; NID:g12514977; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yciC
C:Superfamily: yciC protein

Query Match 46.9%; Score 38; DB 2; Length 247;
Best Local Similarity 60.0%; Pred. No. 74;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSLMSWL 12
| : : : : |
Db 189 LVAPAVLSWL 198

RESULT 34
B64873
probable membrane protein yciC - *Escherichia coli* (strain K-12)
C:Species: *Escherichia coli*
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
A:Accession: B64873; S07798
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64873
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-247 <BLAT>

Query Match	46.9%	Score 38;	DB 1;	Length 267;
Best Local Similarity	57.1%	Pred. No. 79;		
Matches	8;	Conservative	0;	Mismatches 6; Indels 0; Gaps 0;
Qy	1	CELVGPSLMSWLTG 14		
Db	52	CEGKLPSLKIWEYC 65		
RESULT 36				
C95847				
hypothetical transmembrane protein SMD20040 [imported] - Sinorhizobium meliloti				
C;Species:	Sinorhizobium meliloti			
C;Date:	24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004			
C;Accession:	C95847			
R;Finan,	T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Barloy			
P;Finan,	T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Barloy			
A;Title:	The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing			
A;Reference number:	A95842; MUID:21396508; PMID:11481431			
A;Accession:	C95847			
A;Status:	preliminary			
A;Molecule type:	DNA			
A;Residues:	1-286 <KUR>			
A;Cross-references:	UNIPROT:Q92XB9; UNIPARC:UPI00000D4784; GB:AL591985; PIDN:CAO3			
A;Experimental source:	strain 1021, megaplasmid pSymB			
R;Galibert,	F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy			
pela, D.; Chain,	P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher			
L.; Hyman,	R.W.; Jones, T.			
Science	293, 668-672, 2001			
A;Authors:	Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.;			
hebutel, P.; Vandenbol,	M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.;			
A;Title:	The composite genome of the legume symbiont Sinorhizobium meliloti.			
A;Reference number:	A96039; MUID:21368234; PMID:11474104			
A;Contents:	annotation			
C;Genetics:				
A;Gene:	SMB20040			
A;Genome:	plasmid			
Query Match	46.9%	Score 38;	DB 2;	Length 286;
Best Local Similarity	70.0%	Pred. No. 85;		
Matches	7;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
Qy	3	LIVGPSLMSWL 12		
Db	203	LAGPSAASWL 212		
RESULT 37				
T09355				
hypothetical protein F23K16.20 - Arabidopsis thaliana				
C;Species:	Arabidopsis thaliana (mouse-ear cress)			
C;Date:	11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004			
C;Accession:	T09355			
R;Bavan,	M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayev			
submitted to the Protein Sequence Database,	June 1999			
A;Reference number:	Z16652			
A;Accession:	T09355			
A;Molecule type:	DNA			
A;Residues:	1-342 <BEV>			
A;Cross-references:	UNIPROT:Q9SVB8; UNIPARC:UPI000009EFB1; EMBL:AL078620; GSPDB:			
A;Experimental source:	cultivar Columbia; BAC clone F23K16			
C;Genetics:				
A;Gene:	ATSP:F23K16.20			
A;Map position:	4			
A;Introns:	45/1; 106/3; 173/3; 227/3			
Query Match	46.9%	Score 38;	DB 2;	Length 342;
Best Local Similarity	72.7%	Pred. No. 1e+02;		
Matches	8;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
Qy	3	LIVGPSLMSWL 13		

Db 203 LVGPFLLDYMLT 213

RESULT 38
T12379
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Panorpa japonica mitochondrion
C:Species: mitochondrion Panorpa japonica
C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12379
R:Ogai, H.
submitted to the EMBL Data Library, March 1998
A:Description: A genetic evidence showing the identity between form japonica and form kl
A:Reference number: Z17500
A:Accession: T12379
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <OGA>
A:CROSS-references: UNIPROT:O63702; UNIPARC:UPI0000092D34; EMBL:AF056494; NID:g3063638;
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 46.9%; Score 38; DB 2; Length 357;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPFSLMSWL 12
: : : : :
Db 213 VLGGSLMSWL 222

RESULT 39
B29336
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C>Date: 31-Dec-1988 #sequence_revision 22-Jul-1994 #text_change 09-Jul-2004
C:Accession: B29336; B25405; S09373
R:Davidson, E.; Daldal, F.
J. Mol. Biol. 195, 13-24, 1987
A:Title: Primary structure of the bc-1 complex of Rhodospseudomonas capsulata. Nucleotide
A:Reference number: A92938; MUID:88011223; PMID:2821268
A:Accession: B29336
A:Molecule type: DNA
A:Residues: 1-437 <DAV>
A:CROSS-references: UNIPROT:P08502; UNIPARC:UPI000016FE4D; EMBL:X05630; NID:g46093; PIDN
R:Gabbellini, N.; Sebal, W.
Eur. J. Biochem. 154, 569-579, 1986
A:Title: Nucleotide sequence and transcription of the fbc operon from Rhodospseudomonas s
A:Reference number: A91162; MUID:86136096; PMID:3004982
A:Accession: B25405
A:Molecule type: DNA
A:Residues: 1-66, 'ID', 69-280, 'I', 282-437 <GAB>
A:CROSS-references: UNIPARC:UPI000016FE34; EMBL:X03476; NID:g46007; PIDN:CAA27195.1; PID
C:Genetics:
A:Gene: fbcB; petB
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase
F:26-381/Domain: cytochrome b homology <CB6>
F:26-225/Domain: cytochrome b6 homology <CB6>
F:51-67/Domain: transmembrane #status predicted <TM1>
F:96-114/Domain: transmembrane #status predicted <TM2>
F:134-150/Domain: transmembrane #status predicted <TM3>
F:146-193/Domain: periplasmic #status predicted <PER1>
F:195-217/Domain: transmembrane #status predicted <TM4>
F:245-381/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:253-269/Domain: transmembrane #status predicted <TM5>
F:270-329/Domain: periplasmic #status predicted <PER2>
F:330-346/Domain: transmembrane #status predicted <TM6>
F:365-383/Domain: transmembrane #status predicted <TM7>
F:395-411/Domain: transmembrane #status predicted <TM8>

F;97,198/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;111,212/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 46.9%; Score 38; DB 1; Length 437;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSPSLMSWL 12
: : : : :
Db 172 IGPSIQAWL 180

RESULT 40
B69004
hypothetical protein MTH1028 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69004
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69004
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-441 <MTH>
A:CROSS-references: UNIPROT:O27107; UNIPARC:UPI0000062B87; GB:AE000875; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1028

Query Match 46.9%; Score 38; DB 2; Length 441;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELVGPSPSLMSWL 13
: : : : :
Db 194 DVUGPATQSWRT 205

RESULT 41
AG0308
isochorismate synthase (EC 5.4.99.6) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 12-Jun-2003
C:Accession: AG0308
R:Packhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0308
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <KUR>
A:CROSS-references: UNIPARC:UPI00000CDP92A; GB:AL590842; PIDN:CAC91331.1; PID:g15980520;
C:Genetics:
A:Gene: menF
C:Superfamily: isochorismate synthase
C:Keywords: intramolecular transferase; isomerase

Query Match 46.9%; Score 38; DB 2; Length 455;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGSPSLMSWL 13
: : : : :
Db 35 LVGRQLLEWLT 45

RESULT 42

A13367
multidrug resistance protein B [imported] - *Brucella melitensis* (strain 16M)
C/Species: *Brucella melitensis*
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: A13367
R;DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, C.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, P.; Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A/Reference number: AD3252; PMID:11756688
A/Accession: A13367
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-472 <KUR>
A/Cross-references: UNIPROT:Q8VH77; UNIPARC:UPI0000057E89; GB:AE008917; PIDN:AAU52108.1;
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BMEI0927
A/Map position: I

Query Match 46.9%; Score 38; DB 2; Length 472;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LVGPSSLSWLT 13

Db 113 ILGPTLGGWLT 123

RESULT 43

T48630
high affinity nitrate transporter-like protein - *Arabidopsis thaliana*
N/Alternate names: protein T15N1.60
C/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: T48630
R;Bavan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24493
A/Accession: T48630
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-493 <BEV>
A/Cross-references: UNIPROT:Q9LYK2; UNIPARC:UPI0000048B31; EMBL:AL163792
A/Experimental source: cultivar Columbia; BAC clone T15N1
C/Genetics:
A/Map position: 5
A/Introns: 395/2
A/Note: T15N1.60
C/Superfamily: nitrate transporter component

Query Match 46.9%; Score 38; DB 2; Length 493;
Best Local Similarity 41.2%; Pred. No. 1.4e+02;
Matches 7; Conservative 6; Mismatches 0; Indels 4; Gaps 1;

QY 1 CELVGP-----SLMSWLT 13

Db 99 CDLIGPRTSSAILSLT 115

RESULT 44

H83254
probable MFS transporter PA1317 [imported] - *Pseudomonas aeruginosa* (strain PAO1)
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2004
C/Accession: H83254
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho
A/Reference number: A82950; PMID:20437337; PMID:10984043

A/Accession: H83254

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-499 <STO>

A/Cross-references: UNIPROT:Q9HZ88; UNIPARC:UPI000000C5925; GB:AE004737; GB:AE004091; NID

A/Experimental source: strain PAO1

C/Genetics:

A/Gene: PA1317

Query Match 46.9%; Score 38; DB 2; Length 499;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 VGPSLSMSWLT 13

Db 138 IGTGLGGWLT 147

RESULT 45

A70022

multidrug-efflux transporter homolog yusP - *Bacillus subtilis*C/Species: *Bacillus subtilis*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: A70022

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A/Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; PMID:98044033; PMID:9384377
A/Accession: A70022

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-541 <KUN>

A/Cross-references: UNIPROT:O32182; UNIPARC:UPI0000060A25; GB:Z99120; GB:AL009126; NID:G

A/Experimental source: strain 168

C/Genetics:

A/Gene: yusP

C/Superfamily: multidrug-efflux transporter

C/Keywords: antibiotic resistance; transmembrane protein

Query Match 46.9%; Score 38; DB 1; Length 541;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELVGPSLSMSWLT 13

Db 43 DLGGLSMTWLT 54

RESULT 46

T22072

hypothetical protein F41D3.10 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T22072

R;White, S.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19508

A/Accession: T22072

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-594 <WIL>

A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB4; EMBL:Z81537; PIDN:CAB04381.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.10
A;Map position: 1
A;Introns: 80/3; 127/1; 194/2; 250/1; 356/3; 399/1; 437/2; 507/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 594;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CELVGPISLMSWLT 14
Db 400 CEVLPSLVKGMVC 413
::: |||: |

RESULT 47
T22067
hypothetical protein F41D3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22067
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19508
A;Accession: T22067
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-644 <WIL>
A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB3; EMBL:Z81537; PIDN:CAB04376.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.5
A;Map position: 1
A;Introns: 42/2; 130/3; 177/1; 244/2; 300/1; 406/3; 449/1; 487/2; 557/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 644;
Best Local Similarity 42.9%; Pred. No. 1.8e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CELVGPISLMSWLT 14
Db 450 CEVLPSLVKGMVC 463
::: |||: |

RESULT 48
H75485
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75485
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-718 <WHI>
A;Cross-references: UNIPROT:Q9RWG1; UNIPARC:UPI00000C17E3; GB:AE001927; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0707
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0707

Query Match 46.9%; Score 38; DB 2; Length 718;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB4; EMBL:Z81537; PIDN:CAB04381.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.10
A;Map position: 1
A;Introns: 80/3; 127/1; 194/2; 250/1; 356/3; 399/1; 437/2; 507/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 594;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CELVGPISLMSWLT 14
Db 400 CEVLPSLVKGMVC 413
::: |||: |

RESULT 47
T22067
hypothetical protein F41D3.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22067
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19508
A;Accession: T22067
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-644 <WIL>
A;Cross-references: UNIPROT:O45509; UNIPARC:UPI0000179FB3; EMBL:Z81537; PIDN:CAB04376.1;
A;Experimental source: clone F41D3
C;Genetics:
A;Gene: CESP:F41D3.5
A;Map position: 1
A;Introns: 42/2; 130/3; 177/1; 244/2; 300/1; 406/3; 449/1; 487/2; 557/3
C;Superfamily: Caenorhabditis elegans hypothetical protein T26H2.7

Query Match 46.9%; Score 38; DB 2; Length 644;
Best Local Similarity 42.9%; Pred. No. 1.8e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CELVGPISLMSWLT 14
Db 450 CEVLPSLVKGMVC 463
::: |||: |

RESULT 48
H75485
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75485
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75485
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-718 <WHI>
A;Cross-references: UNIPROT:Q9RWG1; UNIPARC:UPI00000C17E3; GB:AE001927; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0707
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0707

Query Match 46.9%; Score 38; DB 2; Length 718;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWLT 13
Db 296 QLGSGLLEWLT 307
::: |||: |

RESULT 49
T04970

probable potassium transport protein KTS - Arabidopsis thaliana
N;Alternate names: protein T16L1.20

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T04970

R;Bevan, M.; Obermaier, B.; Deutschenbaur, S.; Piravandi, E.; Hoheisel, J.; Mewes, H.W.;

submitted to the Protein Sequence Database, November 1998

A;Reference number: Z15393

A;Accession: T04970

A;Molecule type: DNA

A;Residues: 1-846 <BEV>

A;Cross-references: UNIPROT:Q8LPL8; UNIPARC:UPI000009E991; EMBL:AL031394

A;Experimental source: cultivar Columbia; BAC clone T16L1

C;Genetics:

A;Map position: 4

A;Introns: 98/3; 175/1; 268/1; 355/1; 372/3; 409/1; 494/1; 668/2

A;Note: T16L1.20

C;Superfamily: barley probable potassium transport protein HAK1

C;Keywords: ion transport

Query Match 46.9%; Score 38; DB 2; Length 846;
Best Local Similarity 41.7%; Pred. No. 2.4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LVGPISLMSWLT 14
Db 295 VLGPALLWFFC 306
::: |||: |

RESULT 50
B96553

hypothetical protein F5D21.7 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: B96553

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B96553

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-941 <STO>

A;Cross-references: UNIPROT:Q9C8K0; UNIPARC:UPI00000A5EC3; GB:AE005173; NID:G10092355;

C;Genetics:

A;Gene: F5D21.7

A;Map position: 1

Query Match 46.9%; Score 38; DB 2; Length 941;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ELVGPSLMSWL 12
Db 845 KLVGPRDLSQL 855
::: |||: |

Search completed: May 12, 2006, 10:51:51

Job time : 31.0915 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 99.0769 Seconds

(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-67

Perfect score: 81

Sequence: 1 CELVGPLMSLWLC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	61.7	508	2	Q6GNG6_XENLA
2	50	61.7	526	2	Q4V789_XENTR
3	49	60.5	199	1	RNS11_HUMAN
4	49	60.5	199	2	Q5GAN5_HUMAN
5	49	60.5	500	2	Q9VKC2_DROME
6	49	60.5	1510	2	Q7XLY1_ORYSA
7	47	58.0	167	2	Q7XFQ6_ORYSA
8	47	58.0	167	2	Q8W5P2_ORYSA
9	47	58.0	745	2	Q9HE61_NEUCR
10	46	56.8	365	2	O54512_YEREN
11	46	56.8	365	2	Q9Z3C6_YERPE
12	46	56.8	366	2	Q8ZF25_YERPE
13	46	56.8	386	2	Q8D0C4_YERPE
14	46	56.8	452	2	Q5SWU3_CRYNE
15	46	56.8	452	2	Q5KJF5_CRYNE
16	46	56.8	716	2	Q8VXB0_ORYSA
17	46	56.8	778	2	Q652J4_ORYSA
18	46	56.8	1215	2	Q4P7B5_USTMA
19	45	55.6	305	2	Q6ATC7_PROAC
20	45	55.6	328	2	Q6NEV8_CORDI
21	45	55.6	497	2	Q8GAG2_ARTNI
22	45	55.6	524	1	ARHG3_MOUSE
23	45	55.6	526	1	ARHG3_HUMAN
24	45	55.6	526	2	Q5R6F2_PONPY
25	45	55.6	537	2	Q59F00_HUMAN
26	45	55.6	566	2	Q4QQOQ_HUMAN
27	45	55.6	566	2	Q4QQO5_HUMAN
28	45	55.6	734	2	Q52BN4_MAGGR
29	45	55.6	907	2	Q8YZ60_ANASP
30	45	55.6	907	2	Q8YZ60_ANASP
31	44	54.3	481	2	Q5TM27_MACCU

32	44	54.3	481	2	Q8GA10_ARTNI
33	44	54.3	506	2	Q504F6_BRARE
34	44	54.3	836	2	Q8SNA3_MOUSE
35	44	54.3	1029	2	Q5SQ40_HUMAN
36	44	54.3	1077	2	Q5SQ42_HUMAN
37	44	54.3	1078	2	Q5SQ43_HUMAN
38	44	54.3	1096	2	Q6MG49_RAT
39	44	54.3	1098	2	Q9WTN8_RAT
40	44	54.3	1126	2	Q5HYL9_HUMAN
41	44	54.3	1126	2	Q96SA6_HUMAN
42	44	54.3	1126	2	Q9BCN4_HUMAN
43	44	54.3	1132	1	BAT3_HUMAN
44	44	54.3	1132	2	O5STC1_HUMAN
45	44	54.3	1154	2	Q9ZLR2_MOUSE
46	44	54.3	1162	2	Q5SQ39_HUMAN
47	44	54.3	1229	2	Q95874_HUMAN
48	44	54.3	1229	2	O5SQ35_HUMAN
49	43	53.1	108	2	Q4HCW7_9DEIO
50	43	53.1	224	2	Q9Z537_STRCO
51	43	53.1	240	2	Q8TQY4_METAC
52	43	53.1	388	2	Q6D048_ERWCT
53	43	53.1	457	2	Q6W1T9_RHISN
54	43	53.1	476	2	O6I3C3_BACAN
55	43	53.1	487	2	Q95K03_MACFA
56	43	53.1	505	2	Q73ZV0_MYCPA
57	43	53.1	526	2	Q8H0S2_ARATH
58	43	53.1	526	2	Q941F0_ARATH
59	43	53.1	537	2	Q4MNI6_BACCE
60	43	53.1	537	2	Q6HNF1_BACHK
61	43	53.1	537	2	Q73D11_BACCI
62	43	53.1	537	2	Q63FY3_BACCC
63	42.5	52.5	262	1	NASD_KLEOX
64	42.5	52.5	262	1	Q8Y2F7_RALSO
65	42.5	52.5	571	2	Q7SC40_NEUCR
66	42	51.9	74	2	Q95L31_HORSE
67	42	51.9	139	2	Q60903_MOUSE
68	42	51.9	161	2	Q81385_BACCR
69	42	51.9	162	2	Q4MT25_BACCE
70	42	51.9	248	1	Y4UH_RHISN
71	42	51.9	271	2	Q94UH0_9HEMI
72	42	51.9	289	2	Q4T411_AZOVI
73	42	51.9	366	2	Q8FGE7_ECOL6
74	42	51.9	401	2	O00572_HUMAN
75	42	51.9	419	1	PELI2_MOUSE
76	42	51.9	419	2	Q6P4Y5_XENTR
77	42	51.9	420	1	PELI2_HUMAN
78	42	51.9	428	2	Q5VXH9_HUMAN
79	42	51.9	428	2	Q6DHU8_BRARE
80	42	51.9	449	2	Q9FFU8_ARATH
81	42	51.9	487	1	BAT1_HUMAN
82	42	51.9	487	2	Q53FY4_HUMAN
83	42	51.9	497	1	CP26A_HUMAN
84	42	51.9	497	2	Q5VX10_HUMAN
85	42	51.9	517	2	Q7Y5Y4_AEDAE
86	42	51.9	518	2	Q89VM4_BRAJA
87	42	51.9	519	2	Q5DTF7_MOUSE
88	42	51.9	520	2	Q6NCF8_RHOPA
89	42	51.9	539	2	O22252_ARATH
90	42	51.9	542	2	Q6AP84_DPSPS
91	42	51.9	543	2	Q7XMO7_ORYSA
92	42	51.9	549	2	Q8CIG1_MOUSE
93	42	51.9	550	2	Q8R1D1_MOUSE
94	42	51.9	564	2	Q8BJ29_MOUSE
95	42	51.9	621	2	Q03192_YEAST
96	42	51.9	632	2	Q580Y3_9TRYP
97	42	51.9	632	2	Q580Y5_9TRYP
98	42	51.9	686	2	Q07492_9TRYP
99	42	51.9	686	2	Q8WPU7_9TRYP
100	42	51.9	786	2	P91451_CAEEL
101	42	51.9	845	2	Q7XIJ6_ORYSA
102	42	51.9	869	2	YD95_YEAST
103	42	51.9	1005	2	Q75KS9_ORYSA
104	42	51.9	1005	2	Q75KS9_ORYSA

Q8ga10	arthrobacte
Q504f6	brachydanio
Q8sna3	mus musculu
Q5sq40	homo sapien
Q5sq42	homo sapien
Q5sq43	homo sapien
Q6mg49	rattus norv
Q9wt8	rattus norv
Q5hy19	homo sapien
Q96sa6	homo sapien
Q9bcn4	homo sapien
P46379	homo sapien
Q5stc1	homo sapien
Q5stc1	mus musculu
Q5zlr2	homo sapien
Q5sq39	homo sapien
Q95874	homo sapien
Q5sq35	homo sapien
Q4hcm7	deinococcus
Q9z537	streptomyce
Q8tqy4	methanosarc
Q6d048	erwinia car
Q6w1t9	rhizobium s
Q6i3c3	bacillus an
Q95k03	macaca fasc
Q73zv0	mycobacteri
Q8h0s2	arabidopsis
Q941f0	arabidopsis
Q4mni6	bacillus ce
Q6hnf1	bacillus th
Q73d11	bacillus ce
Q63fy3	bacillus ce
P39459	klebsiella
Q8y2f7	raistonia s
Q7sc40	neurospora
Q95l31	equus cabal
Q60903	mus musculu
Q81385	bacillus ce
Q4mt25	bacillus ce
Q53200	rhizobium s
Q94uh0	trialeurode
Q4t411	azotobacter
Q8fge7	escherichia
O00572	homo sapien
Q8bst6	mus musculu
Q6p4y5	xenopus tro
Q9hat8	homo sapien
Q5vxh9	homo sapien
Q6dhj8	brachydanio
Q9ffj8	arabidopsis
P82251	homo sapien
Q53fy4	homo sapien
Q43174	homo sapien
Q5vx10	homo sapien
Q779y4	aedes aegyp
Q89vw4	bradyrhizob
Q5dtf7	mus musculu
Q6ncf8	rhodopseudo
O22252	arabidopsis
Q6ap84	desulfotale
Q7xmo7	oryza sativ
Q8c1g1	mus musculu
Q8r1d1	mus musculu
Q8bj29	mus musculu
Q03192	saccharomyc
Q580y2	trypanosoma
Q580y3	trypanosoma
Q580y5	trypanosoma
Q07492	trypanosoma
Q8wpu7	trypanosoma
P91451	caenorhabdi
Q7xij6	oryza sativ
P39666	saccharomyc
Q75ks9	oryza sativ

105	42	51.9	1128	2	Q850V9_ORYSA	Q850V9 oryza sativ	178	40	49.4	439	2	Q6AB13_PROAC	Q6AB13 propionibac
106	42	51.9	1299	2	Q8H8K3_ORYSA	Q8H8K3 oryza sativ	179	40	49.4	445	1	PELI13_MOUSE	Q8BX6 mus musculu
107	42	51.9	1411	2	Q5KQF3_ORYSA	Q5KQF3 oryza sativ	180	40	49.4	447	2	Q85XK0_MAIZE	Q92X0 zea mays (m
108	42	51.9	1539	2	Q7XP45_ORYSA	Q7XP45 oryza sativ	181	40	49.4	459	2	Q85843_SPHAR	Q92X03 sphingomona
109	42	51.9	1577	2	Q60DR2_ORYSA	Q60DR2 oryza sativ	182	40	49.4	461	2	Q53D15_MACACA	Q53D15 macaca fusc
110	42	51.9	1584	2	Q8LMV8_ORYSA	Q8LMV8 oryza sativ	183	40	49.4	469	1	PELI13_HUMAN	Q82N19 homo sapien
111	42	51.9	1739	2	Q7XXJ0_ORYSA	Q7XXJ0 oryza sativ	184	40	49.4	477	1	SYE_SYNXP	Q7U531 synechococc
112	42	51.9	1975	2	Q84WV2_ORYSA	Q84WV2 oryza sativ	185	40	49.4	519	2	Q8WZS4_ORYSA	Q8WZS4 oryza sativ
113	42	51.9	2011	2	Q8H8S1_ORYSA	Q8H8S1 oryza sativ	186	40	49.4	524	2	Q51VA5_MAGGR	Q51VA5 magnaporth
114	41.5	51.2	187	2	Q7UP94_RHOBA	Q7UP94 rhodospirell	187	40	49.4	537	2	Q81HX2_BACCR	Q81HX2 bacillus ce
115	41	50.6	121	2	Q562B8_RAT	Q562B8 rattus norv	188	40	49.4	541	2	Q59EH4_HUMAN	Q59EH4 homo sapien
116	41	50.6	126	2	Q5HX24_GLUOX	Q5HX24 gluconobact	189	40	49.4	554	1	ZN426_HUMAN	Q3BUV5 homo sapien
117	41	50.6	128	2	Q9GAB0_GRIVA	Q9GAB0 pitar morrh	190	40	49.4	578	2	Q8LMP6_ORYSA	Q8LMP6 oryza sativ
118	41	50.6	144	2	Q62AY7_BURMA	Q62AY7 burkholderi	191	40	49.4	647	2	Q61M45_CAEBR	Q61M45 caenorhabdi
119	41	50.6	263	1	PENK_BOVIN	P01211 bos taurus	192	40	49.4	651	2	O16242_CAEEL	O16242 caenorhabdi
120	41	50.6	285	2	Q8C9D2_MOUSE	Q8C9D2 mus musculu	193	40	49.4	691	2	Q8VY34_MAIZE	Q8VY34 zea mays (m
121	41	50.6	290	2	Q9A9D5_CAUCR	Q9A9D5 caulobacter	194	40	49.4	713	2	Q59GS5_HUMAN	Q59GS5 homo sapien
122	41	50.6	291	2	Q8CBQ3_MOUSE	Q8CBQ3 mus musculu	195	40	49.4	723	2	Q6CSV9_KULUA	Q6CSV9 kluyveromyc
123	41	50.6	361	1	TGT_RICPR	Q9ZCH8 rickettsia	196	40	49.4	809	2	Q94166_ORYSA	Q94166 oryza sativ
124	41	50.6	361	1	TGT_RICRY	Q8WZ6 rickettsia	197	40	49.4	852	2	Q8DGP5_SYNEL	Q8DGP5 synechococc
125	41	50.6	389	1	CHS5_MEDSA	P51078 medicago sa	198	40	49.4	877	2	Q8WZU1_ORYSA	Q8WZU1 oryza sativ
126	41	50.6	405	2	Q5ZKT7_CHICK	Q5ZKT7 gallus gall	199	40	49.4	1010	2	Q8H8I0_ORYSA	Q8H8I0 oryza sativ
127	41	50.6	418	1	PELI11_HUMAN	Q96FA3 homo sapien	200	40	49.4	1048	2	Q8VY36_MAIZE	Q8VY36 zea mays (m
128	41	50.6	418	1	PELI11_MOUSE	Q8C669 mus musculu	201	40	49.4	1063	2	Q9XBJ4_MAIZE	Q9XBJ4 zea mays (m
129	41	50.6	418	2	Q53T26_HUMAN	Q53T26 homo sapien	202	40	49.4	1088	2	O24587_MAIZE	O24587 zea mays (m
130	41	50.6	418	2	Q5SRW7_MOUSE	Q5SRW7 mus musculu	203	40	49.4	1077	1	ADCY4_HUMAN	Q8TFM4 homo sapien
131	41	50.6	418	2	Q7ZXU3_XENLA	Q7ZXU3 xenopus lae	204	40	49.4	1105	2	Q86T27_HUMAN	Q86T27 homo sapien
132	41	50.6	419	2	Q6GQ57_XENLA	Q6GQ57 xenopus lae	205	40	49.4	1156	2	Q7XS40_ORYSA	Q7XS40 oryza sativ
133	41	50.6	428	2	Q5Z129_NOCPA	Q5Z129 nocardia fa	206	40	49.4	1358	2	Q6L3H0_SOLD	Q6L3H0 solanum dem
134	41	50.6	440	2	Q4SB21_TETNG	Q4SB21 tetraodon n	207	40	49.4	1481	2	Q4TC89_TETNG	Q4TC89 tetraodon n
135	41	50.6	471	2	Q8ZX11_PYRAE	Q8ZX11 pyrobaculum	208	40	49.4	1512	2	Q7XBD2_MAIZE	Q7XBD2 zea mays (m
136	41	50.6	487	1	BAT1_MOUSE	Q9QX86 mus musculu	209	40	49.4	1549	2	Q7XWL7_ORYSA	Q7XWL7 oryza sativ
137	41	50.6	487	1	BAT1_RAT	P82252 rattus norv	210	40	49.4	1553	2	Q8S479_MAIZE	Q8S479 zea mays (m
138	41	50.6	487	2	Q4KM04_RAT	Q4KM04 rattus norv	211	40	49.4	1657	2	Q8W0X4_MAIZE	Q8W0X4 zea mays (m
139	41	50.6	494	2	Q4LSG0_9BURK	Q4LSG0 burkholderi	212	40	49.4	1760	2	Q7XBC6_MAIZE	Q7XBC6 zea mays (m
140	41	50.6	503	1	K1173_MOUSE	Q8DL53 mus musculu	213	40	49.4	1778	2	Q9AY98_ORYSA	Q9AY98 oryza sativ
141	41	50.6	504	1	K1173_HUMAN	Q9HCP6 homo sapien	214	40	49.4	1866	2	Q7XH12_ORYSA	Q7XH12 oryza sativ
142	41	50.6	703	2	O16241_CAEEL	O16241 caenorhabdi	215	40	49.4	1866	2	Q8S5R3_ORYSA	Q8S5R3 oryza sativ
143	41	50.6	747	2	Q8CB19_MOUSE	Q8CB19 mus musculu	216	40	49.4	1892	2	Q8H618_MAIZE	Q8H618 zea mays (m
144	41	50.6	747	2	Q7TPX6_MOUSE	Q7TPX6 mus musculu	217	40	49.4	1896	2	Q84MR5_ORYSA	Q84MR5 oryza sativ
145	41	50.6	757	2	Q8VXA8_ORYSA	Q8VXA8 oryza sativ	218	40	49.4	2319	2	Q8H614_MAIZE	Q8H614 zea mays (m
146	41	50.6	760	2	P91227_CAEEL	P91227 caenorhabdi	219	40	49.4	2397	2	Q6Q7Y4_PAPPR	Q6Q7Y4 paramecium
147	41	50.6	867	2	Q7XPL3_ORYSA	Q7XPL3 oryza sativ	220	39.5	48.8	367	2	Q9KW55_PECOC	Q9KW55 pectobacter
148	41	50.6	879	1	MASTL_HUMAN	Q96GX5 homo sapien	221	39.5	48.8	367	2	Q9WXA3_ERMCA	Q9WXA3 erwinia car
149	41	50.6	1209	2	Q7SZF4_BRARE	Q7SZF4 brachydanio	222	39	48.1	82	2	O8L5T4_QUESS	O8L5T4 quercus sub
150	40.5	50.0	291	2	Q93JM9_9BURK	Q93JM9 cupriavidus	223	39	48.1	136	2	Q5UGD3_BETVU	Q5UGD3 beta vulgar
151	40	49.4	68	2	Q6EB12_CAMJE	Q6EB12 campylobact	224	39	48.1	155	2	Q9R880_9GAMM	Q9R880 cycloclasti
152	40	49.4	84	2	Q8Y074_ANASP	Q8Y074 anabaena sp	225	39	48.1	155	2	Q9ZFW4_9GAMM	Q9ZFW4 cycloclasti
153	40	49.4	89	2	Q8N616_HUMAN	Q8N616 homo sapien	226	39	48.1	158	2	Q7QFK5_ANOGA	Q7QFK5 anopheles g
154	40	49.4	104	2	Q5H1U7_HUMAN	Q5H1U7 homo sapien	227	39	48.1	160	2	Q5LU30_SILUPO	Q5LU30 silicibacte
155	40	49.4	109	2	Q9B034_9VIRU	Q9B034 bacterioph	228	39	48.1	161	2	Q6HKV8_BACCH	Q6HKV8 bacillus th
156	40	49.4	129	2	Q4LBS8_SODGL	Q4LBS8 sodalis glo	229	39	48.1	161	2	Q73AQ6_BACCI	Q73AQ6 bacillus ce
157	40	49.4	179	2	Q5RCD5_PONPY	Q5RCD5 pongo pygma	230	39	48.1	161	2	Q63DE6_BACCC	Q63DE6 bacillus ce
158	40	49.4	280	2	Q5H9X5_ORYSA	Q5H9X5 oryza sativ	231	39	48.1	161	2	Q81SM3_BACAN	Q81SM3 bacillus an
159	40	49.4	286	2	Q93VV0_ARATH	Q93VV0 arabidopsis	232	39	48.1	179	2	Q8YIM2_RALSO	Q8YIM2 ralstonia s
160	40	49.4	287	2	Q9SR31_ARATH	Q9SR31 arabidopsis	233	39	48.1	191	2	Q6MKZ6_BDEBA	Q6MKZ6 bdellovibri
161	40	49.4	304	2	Q9RKV0_STRCO	Q9RKV0 streptomyce	234	39	48.1	195	2	O15640_LEIGHA	O15640 leishmania
162	40	49.4	313	2	Q8W0X5_MAIZE	Q8W0X5 zea mays (m	235	39	48.1	207	2	Q5DQH2_ECOLI	Q5DQH2 escherichia
163	40	49.4	351	2	Q76VE6_9CAUD	Q76VE6 bacterioph	236	39	48.1	208	2	Q68665_9GAMM	Q68665 cycloclasti
164	40	49.4	363	2	Q5BK01_RAT	Q5BK01 rattus norv	237	39	48.1	266	2	P74981_YEREN	P74981 yersinia en
165	40	49.4	372	2	O81073_ARATH	O81073 arabidopsis	238	39	48.1	298	2	Q6NF44_CORDI	Q6NF44 corynebacte
166	40	49.4	373	2	Q67985_RHOOP	Q67985 rhodococcus	239	39	48.1	306	2	O81859_ARATH	O81859 arabidopsis
167	40	49.4	374	2	Q850U6_ORYSA	Q850U6 oryza sativ	240	39	48.1	308	2	O81Q46_BACAN	O81Q46 bacillus an
168	40	49.4	378	2	Q9MNX0_PAGLO	Q9MNX0 pagurus lon	241	39	48.1	308	2	Q63AZ7_BACCC	Q63AZ7 bacillus ce
169	40	49.4	379	2	Q55JK6_CRYNE	Q55JK6 cryptococcu	242	39	48.1	324	1	O10R2_HUMAN	O8NGX6 homo sapien
170	40	49.4	379	2	Q5K9Z5_CRYNE	Q5K9Z5 cryptococcu	243	39	48.1	331	2	Q4JU05_CORJK	Q4JU05 corynebacte
171	40	49.4	383	2	Q97A07_THEVO	Q97A07 thermoplas	244	39	48.1	334	2	Q9KWL7_CORGL	Q9KWL7 corynebacte
172	40	49.4	398	2	Q5YBA1_9SPHN	Q5YBA1 sphingomona	245	39	48.1	334	2	Q7BN99_PSEFP	Q7BN99 pseudomonas
173	40	49.4	410	1	ODBA_PSEPH	P09060 pseudomonas	246	39	48.1	334	2	Q88LX3_PSEPK	Q88LX3 pseudomonas
174	40	49.4	410	2	Q88BQ2_PSEPK	Q88BQ2 pseudomonas	247	39	48.1	335	2	O61FS1_HUMAN	O61FS1 homo sapien
175	40	49.4	410	2	Q911M2_PSEPK	Q911M2 pseudomonas	248	39	48.1	356	2	Q6QWG3_ICMPU	Q6QWG3 ictalurid p
176	40	49.4	419	2	Q6ZTM6_HUMAN	Q6ZTM6 homo sapien	249	39	48.1	374	2	Q6AJF4_DESPS	Q6AJF4 desulfocale
177	40	49.4	436	2	Q86VR7_HUMAN	Q86VR7 homo sapien	250	39	48.1	398	2	Q5S5Y5_SHIV1	Q5S5Y5 human immun

251	39	48.1	423	2	Q8AHK4_9HIV1	Q8ahk4 human immun	324	267	1	PENK_HUMAN	P01210 homo sapien
252	39	48.1	429	2	Q7UTX3_RHOB	Q7utx3 rhodopirell	325	267	2	Q6FHC6_HUMAN	Q6fhc6 homo sapien
253	39	48.1	434	2	Q56AF2_9HIV1	Q56af2 human immun	326	267	2	Q6FHE6_HUMAN	Q6fhe6 homo sapien
254	39	48.1	456	2	Q526J5_MAGGR	Q526j5 magnaporthe	327	268	1	PENK_CAVPO	P47969 cavia porce
255	39	48.1	460	2	Q7WUA0_9GAMM	Q7wua0 cycloclasti	328	268	2	Q5ENN1_HETTR	Q5enn1 heterocapsa
256	39	48.1	465	2	Q9LZD8_ARATH	Q9lzd8 arabidopsis	329	275	2	Q66KB3_XENTR	Q66kb3 xenopus tro
257	39	48.1	481	2	Q51UK6_MAGGR	Q51uk6 magnaporthe	330	284	2	Q93UQ6_9SPHN	Q93uq6 agrobacteri
258	39	48.1	487	2	Q9NIR6_RABIT	Q9nir6 oryctolagus	331	286	2	Q99H31_9NUCL	Q99hb3 rhizobium m
259	39	48.1	498	2	Q7UBB1_SYNPX	Q7ubb1 synechococc	332	290	2	Q92XB9_RHIME	Q92xb9 rhizobium m
260	39	48.1	501	2	Q67D59_9HIV1	Q67d59 human immun	333	290	2	Q7VZE1_BORPE	Q7vze1 bordetella
261	39	48.1	506	2	Q741T8_MYCPA	Q741t8 mycobacteri	334	290	2	Q7WGW7_BORBR	Q7wgv7 bordetella
262	39	48.1	510	2	Q4T2X4_TETNG	Q4t2x4 tetraodon n	335	293	2	Q5FRV2_RHOA	Q5frv2 rhodopsuendo
263	39	48.1	517	2	Q9VB75_DROME	Q9vb75 drosophila	336	296	2	Q8DHQ3_SYNEL	Q8dhq3 synechococc
264	39	48.1	529	2	Q4K961_PSEFS	Q4k961 pseudomonas	337	296	2	Q4ID97_GIBZE	Q4id97 gibberella
265	39	48.1	579	2	Q5KDC7_CRYNE	Q5kdc7 cryptococcu	338	314	2	Q87M69_VIBPA	Q87m69 vibrio para
266	39	48.1	585	2	Q8VXA9_ORYSA	Q8vxa9 oryza sativ	339	315	2	Q6CR48_KLUJA	Q6cr48 kluyveromyc
267	39	48.1	595	1	LNT2_LEPIC	Q7zmd8 leptospira	340	318	2	Q5N4R1_SYNP6	Q5n4r1 synechococc
268	39	48.1	595	1	LNT2_LEPIN	Q7zmd8 leptospira	341	318	2	Q5N4R1_SYNP6	Q5n4r1 synechococc
269	39	48.1	596	2	Q55Q40_CRYNE	Q55q40 cryptococcu	342	320	1	OR2TC_HUMAN	Qng77 homo sapien
270	39	48.1	610	2	Q4IKG6_GIBZE	Q4ikg6 gibberella	343	320	1	OR2TC_HUMAN	Qng77 homo sapien
271	39	48.1	630	2	Q55ZM7_CRYNE	Q55zm7 cryptococcu	344	323	2	Q6FJUB_ECOL6	Q6fjfb escherichia
272	39	48.1	630	2	Q5KN20_CRYNE	Q5kn20 cryptococcu	345	323	2	Q6F982_ACIAA	Q6f982 acinetobact
273	39	48.1	637	2	Q4PGT8_USTWA	Q4pgt8 ustilago ma	346	323	2	Q8Q1G3_BRARE	Q8q1g3 brachydanio
274	39	48.1	653	2	Q4QIU9_LEIMA	Q4qiu9 leishmania	347	329	2	Q8Q145_9HIV1	Q8q145 human immun
275	39	48.1	660	2	Q4S0V6_TETNG	Q4s0v6 tetraodon n	348	329	2	Q8Q145_9HIV1	Q8q145 human immun
276	39	48.1	676	2	Q4QJ48_LEIMA	Q4qj48 leishmania	349	332	2	Q8LB62_ARATH	Q8lb62 arabidopsis
277	39	48.1	735	1	DUR3_YEAST	P33413 saccharomyc	350	332	2	Q8RWM7_ARATH	Q8rwm7 arabidopsis
278	39	48.1	737	2	Q98IS7_RHILO	Q98is7 rhizobium l	351	337	2	Q9SVB8_ARATH	Q9svb8 arabidopsis
279	39	48.1	859	2	Q69R18_ORYSA	Q69r18 oryza sativ	352	342	2	Q8Q1G3_BRARE	Q8q1g3 brachydanio
280	39	48.1	944	2	Q82QU4_STRAW	Q82qu4 streptomyce	353	352	2	Q568W4_BRARE	Q568w4 brachydanio
281	39	48.1	987	2	Q7YZM8_CAEEL	Q7yzm8 caenorhabdi	354	352	2	Q568W4_BRARE	Q568w4 brachydanio
282	39	48.1	1003	2	Q7RSV9_PLAYO	Q7rsv9 plasmodium	355	359	2	Q91BY5_9NUCL	Q91by5 helicoverpa
283	39	48.1	1086	2	Q33K83_9CAUD	Q33k83 temperate p	356	361	2	Q8ZUX8_PYRAE	Q8zux8 pyrobaculum
284	39	48.1	1086	2	Q9A0N0_STRPY	Q9a0n0 streptococc	357	362	2	Q35281_MOUSE	Q35281 mus musculu
285	39	48.1	1086	2	Q8NZS4_STRP8	Q8nzs4 streptococc	358	365	2	Q96L07_HUMAN	Q96l07 homo sapien
286	39	48.1	1093	2	Q8K6G1_STRP3	Q8k6g1 streptococc	359	365	2	Q85Q59_9HYST	Q85q59 thrichomys
287	39	48.1	1093	2	Q5XAS7_STRP6	Q5xas7 streptococc	360	370	1	CBNB_ALCEU	P83763 alcaligenes
288	39	48.1	1710	2	Q4RM05_TETNG	Q4rm05 tetraodon n	361	370	1	TFDD_COMAC	Q9rnz9 comamonas a
289	39	48.1	1730	2	Q7NK50_GLOVI	Q7nk50 gloeobacter	362	370	2	Q8GFH9_ALCXX	Q8gf9h alcaligenes
290	39	48.1	1799	2	Q5KQEB_ORYSA	Q5kqeb oryza sativ	363	370	2	Q9L3W6_PSECL	Q9l3w6 pseudomonas
291	39	48.1	1921	2	Q5JNB3_ORYSA	Q5jnb3 oryza sativ	364	372	2	Q7V818_PROMM	Q7v818 prochloroco
292	39	48.1	1939	2	Q5CRH8_CRYPV	Q5crh8 cryptospori	365	374	2	Q85Q47_9HYST	Q85q47 thrichomys
293	39	48.1	1969	2	Q852C7_ORYSA	Q852c7 oryza sativ	366	378	2	Q70EH4_9RODE	Q70eh4 saccostomus
294	39	48.1	2071	2	Q8GZY7_ORYSA	Q8gzy7 oryza sativ	367	378	2	Q70EH4_9RODE	Q70eh4 saccostomus
295	39	48.1	2145	2	Q8H7T1_ORYSA	Q8h7t1 oryza sativ	368	379	1	CYB_PROCY	Q92u07 proechimys
296	39	48.1	2256	2	Q7XPI7_ORYSA	Q7xpi7 oryza sativ	369	379	2	Q65Z02_9MAXI	Q65z02 megalanhus
297	39	48.1	2526	2	Q4WUP7_ASPFU	Q4wup7 aspergillus	370	379	2	Q85Q58_9HYST	Q85q58 thrichomys
298	38.5	47.5	179	2	Q6DIM7_XENTR	Q6dim7 xenopus tro	371	379	2	Q6XVQ5_PROCY	Q6xvq5 proechimys
299	38.5	47.5	509	2	Q50W76_ENTHI	Q50w76 entamoeba h	372	379	2	Q6XVQ5_PROCY	Q6xvq5 proechimys
300	38	46.9	106	2	Q4TTE9_CABRE	Q4tte9 caenorhabdi	373	379	2	Q508J4_9RODE	Q508j4 perognathus
301	38	46.9	111	2	Q54XP7_DICDI	Q54xp7 dictyosteli	374	380	2	Q52CS2_MAGGR	Q52csr2 magnaporthe
302	38	46.9	112	2	P73045_SYNY3	P73045 synechocyst	375	380	2	Q5TMD2_9CHOR	Q5tmd2 epigonichth
303	38	46.9	126	2	Q74JP5_LACJO	Q74jp5 lactobacill	376	381	2	Q954G3_9RODE	Q954g3 sigmodon to
304	38	46.9	128	2	Q8PN10_XANAC	Q8pn10 xanthomonas	377	381	2	Q8HHP2_9RODE	Q8hhp2 sigmodon to
305	38	46.9	133	2	Q7TLB9_9VIRU	Q7tlb9 marine birn	378	381	2	Q8HHP1_9RODE	Q8hhp1 sigmodon to
306	38	46.9	154	2	Q6ZUM8_HUMAN	Q6zum8 homo sapien	379	381	2	Q8HHP0_9RODE	Q8hhp0 sigmodon to
307	38	46.9	166	2	Q96J54_HUMAN	Q96j54 homo sapien	380	381	2	Q8HC70_9RODE	Q8hc70 sigmodon to
308	38	46.9	172	2	Q4TNN5_9SPHN	Q4tnn5 erythrobact	381	381	2	Q5FR07_GLUOX	Q5fr07 gluconobact
309	38	46.9	181	2	Q5K261_GUITH	Q5k261 guillardia	382	392	2	Q8DGS3_9CHICK	Q8dgs3 synechococc
310	38	46.9	219	1	PIGF_HUMAN	Q07326 homo sapien	383	408	2	Q5F493_9CHICK	Q5f493 gallus gall
311	38	46.9	219	1	PIGF_MOUSE	Q09101 mus musculu	384	410	2	Q5F493_9CHICK	Q5f493 gallus gall
312	38	46.9	219	2	Q61B04_HUMAN	Q61b04 homo sapien	385	411	2	Q6P8K5_MOUSE	Q6p8k5 mus musculu
313	38	46.9	219	2	Q8K338_MOUSE	Q8k338 mus musculu	386	418	2	Q5W910_LINUN	Q5w910 lingua ung
314	38	46.9	221	2	Q8M2E3_9SAUR	Q8m2e3 pachydactyl	387	419	2	Q7NYR8_CHRVO	Q7nyr8 chromobacte
315	38	46.9	230	2	Q6XNZ4_9SAUR	Q6xnz4 phelsuma oc	388	435	2	Q51S05_MAGGR	Q51s05 magnaporthe
316	38	46.9	240	2	Q7U9H9_9SNPX	Q7u9h9 synechococc	389	436	1	CYB_RHOCA	P08502 rhodobacter
317	38	46.9	243	2	Q8MLV9_DROME	Q8mlv9 drosophila	390	438	2	Q6A8A8_PROAC	Q6a8a8 propionibac
318	38	46.9	247	1	YCIC_ECO57	Q8cbw7 escherichia	391	441	2	Q27107_METTH	Q27107 methanobac
319	38	46.9	247	1	YCIC_ECOL6	Q8fhw3 escherichia	392	441	2	Q8DBK1_VIBVU	Q8dbk1 vibrio vuln
320	38	46.9	247	1	YCIC_ECOL1	P21365 escherichia	393	446	2	Q98ER5_RHILO	Q98er5 rhizobium l
321	38	46.9	247	1	YCIC_SHIFL	Q83lc9 shigella fl	394	454	2	Q7PT11_ANOGA	Q7pt11 anophales g
322	38	46.9	247	2	Q9MP34_9HYME	Q9mp34 bothriomyrm	395	455	2	Q8ZDN3_YERPE	Q8zdn3 yersinia pe
323	38	46.9	247	2	Q8HZR8_MACMU	Q8hxr8 macaca mula	396	457	2	Q8D0S5_YERPE	Q8d0s5 yersinia pe

P01210	homo sapien
Q6fnc6	homo sapien
Q6fnc6	homo sapien
P47969	cavia porcea
Q39691	heterocapsa
Q66k36	xenopus tro
Q39kug	agrobacteri
Q99h31	helicoverpa
Q92x28	rhizobium m
Q7vze1	bordetella
Q7gw7	bordetella
Q6n2d7	rhodospheu
Q5fv2	gluconobact
Q8d3q3	synecococc
Q4id97	gibberella
Q8r7m6	vibrio para
Q6cr48	kluyveromyc
Q5n4r1	synecococc
Q8ng76	homo sapien
Q8ng77	homo sapien
Q8fj18	escherichia
Q6f982	acinetobact
Q8k106	human immu
Q8k145	human immu
Q8lbc2	arabidopsis
Q8rsw7	arabidopsis
Q8rsw8	arabidopsis
Q801q3	brachydanio
Q568w4	brachydanio
Q63702	panorpa japp
Q91by5	helicoverpa
Q8zxu8	pyrobaculum
Q35281	mus musculu
Q96107	homo sapien
Q8g959	trichomyes
P83763	alcaligenes
Q9rn29	comamonas a
Q8gfn9	alcaligenes
Q913w6	pseudomonas
Q7v818	prochloroco
Q85q47	trichomyes
Q70eh4	saccharomys
Q70eh3	saccharomys
Q92u7	proechimys
Q85z02	megabalanus
Q85q58	trichomyes
Q6xvq6	proechimys
Q6xvq5	proechimys
Q9g0r5	proechimys
Q508j4	perognathus
Q32ec2	magnaporthe
Q5tm2d	epigonichth
Q954q3	sigmondon t
Q8hhp3	sigmondon t
Q8hhp2	sigmondon t
Q8hhp1	sigmondon t
Q8hc70	sigmondon t
Q8f07	gluconobact
Q9dgs8	synecococc
Q5f493	gallus gallu
Q95rk5	mus musculu
Q35w10	lingula ungu
Q7nyr8	chromobact
Q51s05	magnaporthe
P08502	rhodobacter
Q6a8a8	propionibact
Q27107	methanobact
Q8bkb1	vibrio vuln
Q98er5	rhizobium l
Q7pt11	anopheles p
Q8zd3n	yersinia pes
Q8d0e5	yersinia pe

397	38	46.9	457	2	Q669C6 YERPS	Q669C6 yersinia ps	470	37	45.7	158	1	SMG COXBU	Q83a6 coxiella bu
398	38	46.9	464	2	Q77NK0_9GAMA	Q77nk0 rheus monk	471	37	45.7	166	2	Q8SG46_9D1PT	Q8sg46 gasterophil
399	38	46.9	464	2	Q9WR80_9GAMA	Q9wr80 macaca mula	472	37	45.7	175	2	Q9BA42_ANGST	Q9ba42 anopheles s
400	38	46.9	465	2	Q6DHC7_BRARE	Q6dhc7 brachydanio	473	37	45.7	176	2	Q5EB63_RAT	Q5eb63 rattus norv
401	38	46.9	466	2	Q7MIC1_VIBVY	Q7mic1 vibrio vuln	474	37	45.7	182	2	Q702M8_9D1PT	Q702m8 anopheles f
402	38	46.9	472	2	Q8YH77_BRUME	Q8yh77 bruceella me	475	37	45.7	182	2	Q705A4_9D1PT	Q705a4 anopheles j
403	38	46.9	474	2	Q4TD18_TETNG	Q4tdt8 tetraodon n	476	37	45.7	190	2	Q61VC8_CAEBR	Q61vc8 caenorhabdi
404	38	46.9	477	2	Q8PI71_XANAC	Q8pi71 xanthomonas	477	37	45.7	192	2	Q4T815_TETNG	Q4t815 tetraodon n
405	38	46.9	477	2	Q5N951_ORYSA	Q5n951 oryza sativ	478	37	45.7	195	2	Q71S88_LEIDO	Q71s88 leishmania
406	38	46.9	486	2	Q8FC58_ECOL6	Q8fc58 escherichia	479	37	45.7	195	2	Q71S89_LEICH	Q71s89 leishmania
407	38	46.9	489	2	Q9LYK2_ARATH	Q9lyk2 arabidopsis	480	37	45.7	203	2	Q6TLC5_APIME	Q6tlc5 apis mellif
408	38	46.9	493	2	Q4LUX4_9BURK	Q4lux4 burkholderi	481	37	45.7	208	2	Q71S91_LEIDO	Q71s91 leishmania
409	38	46.9	499	2	Q9HZ88_PSEAE	Q9hz88 pseudomonas	482	37	45.7	208	2	Q71S92_LEICH	Q71s92 leishmania
410	38	46.9	504	2	Q4SD74_TETNG	Q4sd74 tetraodon n	483	37	45.7	220	2	Q85NU0_9SAUR	Q85nu0 pachydaetyl
411	38	46.9	507	2	Q4KFI8_PSEPS	Q4kfi8 pseudomonas	484	37	45.7	221	2	Q8M2D8_9SAUR	Q8m2d8 pachydaetyl
412	38	46.9	511	2	Q983P8_RHILO	Q983p8 rhizobium l	485	37	45.7	225	2	Q41YZ6_AZOVI	Q41yz6 azotobacter
413	38	46.9	513	2	Q8GON0_BRUSU	Q8gon0 bruceella su	486	37	45.7	225	2	Q88EG2_PSEPK	Q88eg2 pseudomonas
414	38	46.9	516	1	LNT CHRVO	Q920e1 mus musculu	487	37	45.7	239	2	Q7QWJ0_GIALA	Q7qwj0 giardia lam
415	38	46.9	516	2	Q94JG1_ORYSA	Q94jg1 oryza sativ	488	37	45.7	245	2	Q9LH11_ARATH	Q9lh11 arabidopsis
416	38	46.9	516	2	Q88LP8_PSEPK	Q88lp8 pseudomonas	489	37	45.7	248	2	Q9HWC1_PSEAE	Q9hwc1 pseudomonas
417	38	46.9	517	2	Q74D39_GEOSL	Q74d39 geobacter s	490	37	45.7	250	2	Q53U42_LYCES	Q53u42 lycopersico
418	38	46.9	526	2	Q7S3U3_NEUCR	Q7s3u3 neurospora	491	37	45.7	250	2	Q8W1A5_PETHY	Q8w1a5 petunia hyb
419	38	46.9	529	2	Q9ZOB1_NOUSE	Q9zob1 mus musculu	492	37	45.7	250	2	Q8SM37_SPOST	Q8sm37 sporobolus
420	38	46.9	530	2	Q8TC57_HUMAN	Q8tc57 homo sapien	493	37	45.7	253	2	Q7PMK2_ANOGA	Q7pmk2 anopheles g
421	38	46.9	533	2	Q7F952_ORYSA	Q7f952 oryza sativ	494	37	45.7	257	2	Q5PX07_HCMV	Q5px07 human cytom
422	38	46.9	533	2	Q5E6H2_VIBP1	Q5e6h2 vibrio fisc	495	37	45.7	260	2	Q49657_MYCLE	Q49657 mycobacteri
423	38	46.9	539	2	Q8EJH2_SHEON	Q8ejh2 shewanella	496	37	45.7	266	1	HMUU YERPE	Q56993 yersinia pe
424	38	46.9	541	2	Q321B2_BACSU	Q321b2 bacillus su	497	37	45.7	266	2	Q623V0_CAEBR	Q623v0 caenorhabdi
425	38	46.9	549	2	Q6UBA5_HUMAN	Q6uba5 homo sapien	498	37	45.7	266	2	Q66FK0_YERPS	Q66fk0 yersinia ps
426	38	46.9	550	2	Q4RQ38_TETNG	Q4rq38 tetraodon n	499	37	45.7	272	2	Q8LAQ8_ARATH	Q8laq8 arabidopsis
427	38	46.9	580	2	Q8IVG3_HUMAN	Q8ivg3 homo sapien	500	37	45.7	274	2	Q984A1_RHILO	Q984a1 rhizobium l
428	38	46.9	645	2	Q97448_GIALA	Q97448 giardia lam	501	37	45.7	284	1	XTH13_ARATH	Q9fk18 arabidopsis
429	38	46.9	649	2	Q61M52_CAEBR	Q61m52 caenorhabdi	502	37	45.7	291	2	Q58PQ7_9POT	Q58pq7 uncultured
430	38	46.9	668	2	Q965R5_CABEL	Q965r5 caenorhabdi	503	37	45.7	295	2	Q6UKY2_9BURK	Q6uky2 burkholderi
431	38	46.9	673	2	Q45509_CABEL	Q45509 caenorhabdi	504	37	45.7	297	1	YCBU ECOLI	P75845 escherichia
432	38	46.9	684	2	Q61AW5_CAEBR	Q61aw5 caenorhabdi	505	37	45.7	297	2	Q57R08_SALCH	Q57r08 salmonella
433	38	46.9	718	2	Q9RWG1_DEIRA	Q9rwg1 deinococcus	506	37	45.7	297	2	Q5PGF6_SALPA	Q5pgf6 salmonella
434	38	46.9	737	2	Q522D6_MAGGR	Q522d6 magnaporthe	507	37	45.7	297	2	Q8FJA8_ECOL6	Q8fja8 escherichia
435	38	46.9	790	2	Q96MR9_HUMAN	Q96mr9 homo sapien	508	37	45.7	297	2	Q827Z9_SALTI	Q827z9 salmonella
436	38	46.9	796	2	Q5AM18_CANAL	Q5am18 candida alb	509	37	45.7	297	2	Q8ZQB9_SALTY	Q8zqb9 salmonella
437	38	46.9	799	1	SYFB PARUW	Q6me6 parachlamyd	510	37	45.7	299	2	Q6UKY1_9BURK	Q6uky1 burkholderi
438	38	46.9	854	1	DRL5_ARATH	Q8c8k arabidopsis	511	37	45.7	299	2	Q9PVL6_CHICK	Q9pvl6 gallus gall
439	38	46.9	855	1	POT13_ARATH	Q8lp18 arabidopsis	512	37	45.7	301	2	Q63JA5_BURPS	Q63ja5 burkholderi
440	38	46.9	869	2	Q5L6V5_CHLAB	Q5l6v5 chlamydophi	513	37	45.7	307	2	Q6UKX9_9BURK	Q6ukx9 burkholderi
441	38	46.9	1077	2	Q7ZZA5_BRARE	Q7zza5 brachydanio	514	37	45.7	307	2	Q6UKY0_9BURK	Q6uky0 burkholderi
442	38	46.9	1110	2	Q8HY22_PIG	Q8hy22 sus scrofa	515	37	45.7	310	1	ORAKE HUMAN	Q8ngd5 homo sapien
443	38	46.9	1116	2	Q07898_HUMAN	Q8hy22 sus scrofa	516	37	45.7	316	2	Q4K7V4_PSEPS	Q4k7v4 pseudomonas
444	38	46.9	1149	2	Q07901_HUMAN	Q07901 homo sapien	517	37	45.7	319	1	MRGRD RAT	Q7tn41 rattus norv
445	38	46.9	1151	2	Q07899_HUMAN	Q07899 homo sapien	518	37	45.7	330	2	Q53RL9_ORYSA	Q53rl9 oryza sativ
446	38	46.9	1156	2	Q86VB7_HUMAN	Q86vb7 homo sapien	519	37	45.7	333	2	Q926G3_RHIME	Q926g3 rhizobium m
447	38	46.9	1156	2	Q07900_HUMAN	Q07900 homo sapien	520	37	45.7	338	2	Q825J4_STREPTOMYCE	Q825j4 streptomyce
448	38	46.9	1337	2	Q4PA65_USTWA	Q4pa65 ustilago ma	521	37	45.7	341	2	Q85700_STRLI	Q85700 streptomyce
449	38	46.9	1410	2	Q51L78_MAGGR	Q51l78 magnaporthe	522	37	45.7	341	2	Q7AKR4_STRCO	Q7akr4 streptomyce
450	38	46.9	1545	1	MRP2_HUMAN	Q22887 homo sapien	523	37	45.7	344	2	Q5AHZ3_CANAL	Q5ahz3 candida alb
451	38	46.9	1545	2	Q5T2E1_HUMAN	Q5t2b1 homo sapien	524	37	45.7	371	2	Q9LDP1_ARATH	Q9ldp1 arabidopsis
452	38	46.9	3165	2	Q5AR63_EMENI	Q5ar63 aspergillus	525	37	45.7	377	2	Q82PK5_STEAM	Q82pk5 streptomyce
453	38	46.9	3618	2	Q4JTA1_CORUK	Q4jta1 corynebacte	526	37	45.7	378	2	Q911Z3_PSEAE	Q911z3 pseudomonas
454	37.5	46.3	336	2	Q6APK5_DESPS	Q6apk5 desulfotale	527	37	45.7	378	2	Q7U759_SYNPX	Q7u759 synecococc
455	37.5	46.3	509	2	Q8GM15_STRGL	Q8gm15 streptomyce	528	37	45.7	380	2	Q66SY3_9SAUR	Q66sy3 rhineura fl
456	37.5	46.3	571	1	PRIM DEIRA	Q9rw5 deinococcus	529	37	45.7	387	2	Q9Y7Y4_SCHPO	Q9y7y4 schizosacch
457	37.5	46.3	627	2	Q4T6E2_TETNG	Q4t6e2 tetraodon n	530	37	45.7	389	2	Q5FHU7_LACAC	Q5fhu7 lactobacill
458	37	45.7	80	2	Q84231_ORYSA	Q84231 oryza sativ	531	37	45.7	390	2	Q90159_HUMIM	Q90159 human immun
459	37	45.7	82	2	Q711O2_LACDL	Q711q2 lactobacill	532	37	45.7	391	2	Q4JOD5_AZOVI	Q4jod5 azotobacter
460	37	45.7	86	2	Q98N14_RHILO	Q98n14 rhizobium l	533	37	45.7	395	2	Q5SRJ5_HUMAN	Q5srj5 homo sapien
461	37	45.7	92	2	Q76YD3_9CAUD	Q76yd3 bacterioph	534	37	45.7	401	1	MDTH CANBF	Q7vqx5 candidatus
462	37	45.7	99	2	Q8ANJ3_9HIV1	Q8anj3 human immun	535	37	45.7	401	2	Q901U9_9HIV1	Q901u9 human immun
463	37	45.7	121	2	Q87KN7_VIBPA	Q87kn7 vibrio para	536	37	45.7	403	2	Q88NU2_PSEPK	Q88nu2 pseudomonas
464	37	45.7	123	2	Q4TRR4_SPHEN	Q4trr4 erythrobact	537	37	45.7	410	2	Q4LN26_9BURK	Q4ln26 burkholderi
465	37	45.7	128	2	Q9GAB2_MERCA	Q9gab2 mercuraria	538	37	45.7	410	2	Q5WTU9_LEGPL	Q5wtu9 legionella
466	37	45.7	134	2	Q9MC43_BPD3	Q9mc43 bacterioph	539	37	45.7	410	2	Q5X223_LEGPA	Q5x223 legionella
467	37	45.7	147	2	Q6ZRM0_HUMAN	Q6zrm0 homo sapien	540	37	45.7	410	2	Q5ZSL6_LEGPH	Q5zsl6 legionella
468	37	45.7	150	2	Q8TS13_METAC	Q8ts13 methanosarc	541	37	45.7	410	2	Q63HZ6_BURPS	Q63hz6 burkholderi
469	37	45.7	150	2	Q5MZ14_SYNP6	Q5mz14 synecococc	542	37	45.7	410	2	Q629V1_BURMA	Q629v1 burkholderi

543	37	45.7	411	2	Q4KDP2_PSEF5	Q4kdp2 pseudomonas	616	37	45.7	668	2	Q68D67_HUMAN	Q68d67 homo sapien
544	37	45.7	415	2	Q7VSR9_PROMM	Q7v5r-9 prochloroco	617	37	45.7	672	2	Q4PIN3_USTMA	Q4pin3 ustlago ma
545	37	45.7	423	2	Q8AQH8_SHIV1	Q8ah8 human immun	618	37	45.7	676	2	Q4NKK3_9MICC	Q4nkz3 arthrobacte
546	37	45.7	426	2	Q6SGV9_PBACT	Q6sgv9 uncultured	619	37	45.7	678	2	Q8MT15_LEITA	Q8mt15 leishmania
547	37	45.7	427	2	Q5WYR1_LEGPL	Q5wyr1 legionella	620	37	45.7	683	2	Q973N2_SULTO	Q973n2 sulfolobus
548	37	45.7	427	2	Q5X7B3_LEGPA	Q5x7b3 legionella	621	37	45.7	684	2	Q9UAC1_LEIDO	Q9uac1 leishmania
549	37	45.7	427	2	Q5ZXU3_LEGPH	Q5zxu3 legionella	622	37	45.7	691	2	Q9UAB7_LEIDO	Q9uab7 leishmania
550	37	45.7	433	1	TUB4_AGRV1	Q44470 agrobacteri	623	37	45.7	691	2	Q4QH10_LEIMA	Q4qh10 leishmania
551	37	45.7	437	2	Q9WYG8_ALCPA	Q9wxg8 alcaligenes	624	37	45.7	698	2	Q9UAC0_LEIDO	Q9uac0 leishmania
552	37	45.7	449	1	TUB3_AGRV1	P70786 agrobacteri	625	37	45.7	700	2	Q9UAB9_LEIDO	Q9uab9 leishmania
553	37	45.7	449	1	Q5ASV3_EMENI	Q5asv3 aspergillus	626	37	45.7	724	2	Q6ESU3_ORYSA	Q6esu3 oryza sativ
554	37	45.7	450	2	Q87MA1_VIBPA	Q87ma1 vibrio para	627	37	45.7	724	2	Q4TCB2_TETNG	Q4tc82 tetraodon n
555	37	45.7	451	2	Q80503_ARATH	Q80503 arabidopsis	628	37	45.7	795	2	Q7PUS2_ANOGA	Q7pus2 anopheles g
556	37	45.7	451	2	Q94A95_ARATH	Q94a95 arabidopsis	629	37	45.7	819	2	Q7PUS2_ANOGA	Q7pus2 anopheles g
557	37	45.7	451	2	Q6I7X8_LEPDU	Q6i7x8 leptotyphlo	630	37	45.7	841	2	Q7Q797_GIALA	Q7qtc97 giardia lam
558	37	45.7	452	2	Q939X8_PSEU	Q939x8 ancycolatops	631	37	45.7	848	1	DRL32_ARATH	Q9f91 arabidopsis
559	37	45.7	454	2	Q98134_HVH8	Q98134 human herpe	632	37	45.7	864	1	XYLT2_RAT	Q9ep10 rattus norv
560	37	45.7	454	2	P88921_HVH8	P88921 human herpe	633	37	45.7	865	1	XYLT2_CANPA	Q5qg50 canis famil
561	37	45.7	457	2	Q8N903_HUMAN	Q8n903 homo sapien	634	37	45.7	865	1	XYLT2_BOVIN	Q5qg49 bos taurus
562	37	45.7	457	2	Q6E7D0_GUROC	Q6e7d0 oikopleura	635	37	45.7	867	1	XYLT2_BOVIN	Q01526 caenorhabdi
563	37	45.7	457	2	Q7Q4U5_ANOGA	Q7q4u5 anopheles g	636	37	45.7	867	1	XYLT2_BOVIN	Q01526 caenorhabdi
564	37	45.7	465	2	Q6FFG2_ACIAD	Q6ffg2 acinetobact	637	37	45.7	868	2	O01526_CAEBL	Q583Q4 trypanosoma
565	37	45.7	470	2	Q8CXN1_OCEIH	Q8cxn1 oceanobacil	638	37	45.7	880	2	Q583Q4_9TRYP	Q583Q4 trypanosoma
566	37	45.7	473	2	Q4UX91_XANCP	Q4ux91 xanthomonas	639	37	45.7	920	2	Q94165_ORYSA	Q94165 oryza sativ
567	37	45.7	473	2	Q8P6X0_XANCP	Q8p6x0 xanthomonas	640	37	45.7	921	2	Q4WVD1_ASPPU	Q4wvd1 aspergillus
568	37	45.7	476	2	Q8XQZ9_RALSO	Q8xqz9 ralstonia s	641	37	45.7	955	2	Q5JNB7_ORYSA	Q5jnb7 oryza sativ
569	37	45.7	478	2	Q4HB58_9DEIO	Q4hb58 deinococcus	642	37	45.7	955	2	Q7Z291_CAEBL	Q7z291 caenorhabdi
570	37	45.7	479	2	Q62HJ7_BURMA	Q62hj7 burkholderi	643	37	45.7	986	2	Q5VS00_ORYSA	Q5vs00 oryza sativ
571	37	45.7	479	2	Q63WK3_BURPS	Q63wk3 burkholderi	644	37	45.7	1001	2	Q7KVS9_DROME	Q7kvs9 drosophila
572	37	45.7	481	1	TRI10_HUMAN	Q9udy6 homo sapien	645	37	45.7	1020	2	Q8IU50_CAEBL	Q8iu50 caenorhabdi
573	37	45.7	481	1	TRI10_PANTR	Q7yr32 pan troglod	646	37	45.7	1020	2	Q19204_CAEBL	Q19204 caenorhabdi
574	37	45.7	481	2	Q5SRK8_HUMAN	Q5srk8 homo sapien	647	37	45.7	1035	2	Q94158_ORYSA	Q94158 oryza sativ
575	37	45.7	481	2	Q5TM54_MACMU	Q5tm54 macaca mula	648	37	45.7	1035	2	Q6CDV8_YARLI	Q6cdv8 yarrowia li
576	37	45.7	481	2	Q4K617_PSEF5	Q4k617 pseudomonas	649	37	45.7	1035	2	Q6CDV8_YARLI	Q6cdv8 yarrowia li
577	37	45.7	484	2	Q7Q1J2_ANOGA	Q7q1j2 anopheles g	650	37	45.7	1096	2	Q4JFH1_CAEBL	Q4jfh1 caenorhabdi
578	37	45.7	500	2	Q9GMO0_BRAJA	Q9gmo0 bradyrhizob	651	37	45.7	1121	2	Q7UPY4_RHOBA	Q7upy4 rhodopirell
579	37	45.7	501	2	Q6AQF7_DESPS	Q6aqf7 desulfotale	652	37	45.7	1128	2	Q6MFB7_PARUW	Q6mfb7 parachlamyd
580	37	45.7	508	2	Q8IMC1_DROME	Q8imc1 dirosophila	653	37	45.7	1195	2	Q6MFB7_PARUW	Q6mfb7 parachlamyd
581	37	45.7	513	2	Q6CJF1_KUJLA	Q6cjp1 kluveromyc	654	37	45.7	1240	2	Q4FYP3_LEIMA	Q4fyp3 leishmania
582	37	45.7	513	2	Q615H8_CAEBR	Q615h8 caenorhabdi	655	37	45.7	1275	2	Q9SXF0_ARATH	Q9sxf0 arabidopsis
583	37	45.7	517	2	Q9V8D9_DROME	Q9v8d9 drosophila	656	37	45.7	1292	2	LIFS2_ASHGO	Q9sxf0 arabidopsis
584	37	45.7	517	2	Q8H595_ORYSA	Q8h595 oryza sativ	657	37	45.7	1309	2	Q9SAI7_ARATH	Q9sa17 arabidopsis
585	37	45.7	527	2	Q88L68_PSEPK	Q88l68 pseudomonas	658	37	45.7	1385	1	Q75G45_ORYSA	Q75g45 oryza sativ
586	37	45.7	538	2	Q8KV00_GROKA	Q8kv00 geobacillus	659	37	45.7	1415	2	Q94KV0_ARATH	Q94kv0 arabidopsis
587	37	45.7	540	2	Q5ZJMO_CHICK	Q5zjmo gallus gall	660	37	45.7	1431	2	Q94KV0_ARATH	Q94kv0 arabidopsis
588	37	45.7	545	2	Q553M5_DICDI	Q553m5 dictyosteli	661	37	45.7	1453	2	Q94IU9_ARATH	Q94iu9 arabidopsis
589	37	45.7	545	2	Q7RUH6_NEUCR	Q7ruh6 neurospora	662	37	45.7	1466	2	Q9FJV3_ARATH	Q9fjv3 arabidopsis
590	37	45.7	546	2	Q4LH51_9BURK	Q4lh51 burkholderi	663	37	45.7	1475	2	Q4MYQ1_THEPA	Q4myq1 theileria p
591	37	45.7	549	2	Q87BE1_XYLFY	Q87be1 xylella fae	664	37	45.7	1732	2	Q8YVY3_ANASP	Q8yvy3 anabaena ap
592	37	45.7	552	2	Q8KU74_ENTFA	Q8ku74 enterococcc	665	37	45.7	1796	2	Q8YVY3_ANASP	Q8yvy3 anabaena ap
593	37	45.7	558	2	Q8KU74_ENTFA	Q8ku74 streptococcc	666	37	45.7	2491	2	Q5TFF4_HUMAN	Q5tff4 homo sapien
594	37	45.7	561	2	Q8K8R4_STRP3	Q8k8r4 streptococcc	667	37	45.7	2566	2	Q5TFF4_HUMAN	Q5tff4 homo sapien
595	37	45.7	564	2	Q55HC1_CRYNE	Q55hc1 cryptococcc	668	37	45.7	2567	1	MY18B_HUMAN	Q9uag5 homo sapien
596	37	45.7	566	2	Q7VAE3_PROMA	Q7vae3 prochloroco	669	36.5	45.1	3856	1	ATM_ARATH	Q9uag5 homo sapien
597	37	45.7	575	2	Q879N8_STRP3	Q879n8 streptococcc	670	36.5	45.1	60	2	Q7QUN1_GIALA	Q7qun1 giardia lam
598	37	45.7	575	2	Q5XE24_STRP6	Q5xe24 streptococcc	671	36.5	45.1	179	2	Q5M783_XENLA	Q5m783 xenopus lae
599	37	45.7	575	2	Q9A1M9_STRPY	Q9a1m9 streptococcc	672	36.5	45.1	197	2	Q7UDX8_RHOBA	Q7udx8 rhodopirell
600	37	45.7	575	2	Q7NJQ7_GLOVI	Q7njq7 gloeobacter	673	36.5	45.1	229	1	RPE_CHLPN	Q92829 chlamydia p
601	37	45.7	584	2	Q4SRM1_TETNG	Q4srml tetraodon n	674	36.5	45.1	469	2	Q4QC26_LEIMA	Q4qc26 leishmania
602	37	45.7	585	2	Q4SRM1_TETNG	Q4srml tetraodon n	675	36.5	45.1	487	2	Q75IF2_ORYSA	Q75if2 oryza sativ
603	37	45.7	595	2	Q5ZMB3_CHICK	Q5zmb3 gallus gall	676	36.5	45.1	555	1	GPC6_HUMAN	Q9y625 homo sapien
604	37	45.7	601	2	Q5G5D5_MOUSE	Q5g5d5 mus musculu	677	36.5	45.1	555	2	Q96SG8_HUMAN	Q96sg8 homo sapien
605	37	45.7	617	2	Q5G5D5_MOUSE	Q5g5d5 mus musculu	678	36.5	45.1	555	2	Q5RE54_PONPY	Q5re54 pongo pygma
606	37	45.7	618	2	Q6NND1_DROME	Q6nnd1 rattus norv	679	36.5	45.1	555	2	Q6FKZ7_CANGA	Q6fkz7 candida gla
607	37	45.7	636	2	Q9V4E7_DROME	Q9v4e7 drosophila	680	36.5	45.1	1096	2	Q7QHH1_ANOGA	Q7qhh1 anopheles g
608	37	45.7	642	2	Q920Q8_MOUSE	Q920q8 mus musculu	681	36.5	45.1	1137	2	Q7QHH1_ANOGA	Q7qhh1 anopheles g
609	37	45.7	642	2	Q99KN0_MOUSE	Q99kn0 mus musculu	682	36.5	45.1	1500	2	Q6R7Y0_PONPY	Q6r7y0 pongo pygma
610	37	45.7	644	2	Q62Q34_MOUSE	Q62q34 mus musculu	683	36.5	45.1	1549	2	Q6PGN0_MOUSE	Q6pgn0 mus musculu
611	37	45.7	654	2	Q4Q5A1_LEIMA	Q4q5a1 leishmania	684	36.5	45.1	1560	2	Q8CGM1_MOUSE	Q8cgm1 mus musculu
612	37	45.7	657	1	KUPI_BRAJA	Q89nm6 bradyrhizob	685	36.5	45.1	1572	1	BA12_HUMAN	Q60241 homo sapien
613	37	45.7	664	2	Q8G4A4_BIFLO	Q8g4a4 bifidobacte	686	36	44.4	1573	2	Q8NGW8_HUMAN	Q8ngw8 homo sapien
614	37	45.7	664	2	Q8G4A4_BIFLO	Q8g4a4 bifidobacte	687	36	44.4	1584	2	Q5TGK0_HUMAN	Q5tgk0 homo sapien
615	37	45.7	665	2	Q4QH12_LEIMA	Q4qh12 leishmania	688	36	44.4	57	2	Q87WAS_PSESM	Q87was pseudomonas
										77	2	Q5B808_SCHJA	Q5b808 schistosoma
										77	2	Q7YUY7_9TRYP	Q7yuy7 trypanosoma

689	36	44.4	77	2	Q5PAK1_ANAMM	O5pak1 anaplasma m	762	36	44.4	267	2	Q8KG89_CHLTE	Q8kg89 chlorobium
690	36	44.4	77	2	O5P9H0_ANAMM	O5p9h0 anaplasma m	763	36	44.4	269	2	Q6NHR3_CORDI	Q6nhr3 corynebacte
691	36	44.4	77	2	O5PBV7_ANAMM	O5pbv7 anaplasma m	764	36	44.4	269	2	Q67MB9_SYMTH	Q67mb9 symbiobacte
692	36	44.4	78	2	Q7WLR9_BORBR	Q7wlr9 bordetella	765	36	44.4	273	2	Q9DBH4_MOUSE	Q9dbh4 mus musculus
693	36	44.4	83	2	Q82W04_NITEU	Q82w04 nitrosomona	766	36	44.4	277	2	Q53K02_ORYSA	Q53k02 oryza sativ
694	36	44.4	89	2	Q83TD1_NITEU	Q83td1 anaplasma m	767	36	44.4	277	2	Q811F9_MOUSE	Q811f9 mus musculus
695	36	44.4	89	2	Q84H53_ANAMA	Q84h53 anaplasma m	768	36	44.4	290	2	Q82BH7_STRAW	Q82bh7 streptomyc
696	36	44.4	89	2	O5P9V5_ANAMM	O5p9v5 anaplasma m	769	36	44.4	295	2	O03320_9SAUR	O03320 elseyia lati
697	36	44.4	91	2	Q84HV1_ANAMA	Q84hv1 anaplasma m	770	36	44.4	297	2	O03335_9SAUR	O03335 pelusios wi
698	36	44.4	99	2	Q9J079_9HIVI	Q9j079 human immun	771	36	44.4	298	2	Q6P519_HUMAN	Q6p519 homo sapien
699	36	44.4	102	2	Q4G007_HUMAN	Q4g007 homo sapien	772	36	44.4	300	2	O51RR3_MAGOR	O51rr3 magnaporthe
700	36	44.4	110	2	Q05841_YEAST	Q05841 saccharomyc	773	36	44.4	300	2	Q628N2_CASBR	Q628n2 caenorhabdi
701	36	44.4	118	2	O60912_MOUSE	O60912 mus musculus	774	36	44.4	301	2	Q8WY00_HUMAN	Q8wy00 homo sapien
702	36	44.4	128	2	Q7XV70_ORYSA	Q7xv70 oryza sativ	775	36	44.4	301	2	Q6AGK6_LEIXX	Q6agk6 leifsonia x
703	36	44.4	137	1	CRCB1_BACCI	C61386 bacillus ce	776	36	44.4	302	2	Q8C790_MOUSE	Q8c790 mus musculus
704	36	44.4	145	2	Q89DP5_BRAJA	Q89dp5 bradyrhizob	777	36	44.4	302	2	Q55VB5_CRYNE	Q55vb5 cryptococcu
705	36	44.4	147	2	Q865E3_CANFA	Q865e3 canis famill	778	36	44.4	302	2	O5KKX0_CRYNE	O5kkx0 cryptococcu
706	36	44.4	149	2	Q896D4_CLOTE	Q896d4 clostridium	779	36	44.4	304	2	Q821K7_SALTI	Q821k7 salmonella
707	36	44.4	166	1	HA22G_ARATH	O91r09 arabidopsis	780	36	44.4	305	2	Q5ZD45_ORYSA	Q5zd45 oryza sativ
708	36	44.4	173	2	Q7NP13_GLOVI	Q7np13 gloeobacter	781	36	44.4	308	1	GSUB_PROMA	Q7tub0 prochloroco
709	36	44.4	174	2	O5YZH1_NOCFA	O5yzh1 nocardia fa	782	36	44.4	309	2	Q8NI49_HUMAN	Q8ni49 homo sapien
710	36	44.4	175	2	Q9B496_9DIPT	Q9b496 anopheles b	783	36	44.4	309	2	Q9JKT8_RAT	Q9jkt8 rattus norv
711	36	44.4	175	2	Q9B497_9DIPT	Q9b497 anopheles c	784	36	44.4	313	1	COFG_SYNY3	P73191 synchocyst
712	36	44.4	175	2	Q9BA45_ANOAL	Q9ba45 anopheles a	785	36	44.4	315	2	Q34185_CEPNE	Q34185 cepaea nemo
713	36	44.4	175	2	Q9BA46_ANOAB	Q9ba46 anopheles a	786	36	44.4	318	2	O5J343_ORYSA	Q5j343 oryza sativ
714	36	44.4	182	2	Q702M9_9DIPT	Q702m9 anopheles a	787	36	44.4	319	2	Q5UF78_9PROT	Q5uf78 uncultured
715	36	44.4	182	2	Q702N0_9DIPT	Q702n0 anopheles a	788	36	44.4	320	2	Q4UZV7_XANCP	Q4uzv7 xanthomonas
716	36	44.4	182	2	Q702N2_9DIPT	Q702n2 anopheles a	789	36	44.4	320	2	Q8PDL6_XANCP	Q8pdl6 xanthomonas
717	36	44.4	182	2	Q702N3_9DIPT	Q702n3 anopheles a	790	36	44.4	321	2	Q4KEA0_PSEF5	Q4kea0 pseudomonas
718	36	44.4	182	2	Q702N5_9DIPT	Q702n5 anopheles a	791	36	44.4	323	2	Q7U936_SYNPX	Q7u936 synchococc
719	36	44.4	192	2	O6N4D2_RHOPA	O6n4d2 rhodopseudo	792	36	44.4	324	1	TTP_BOVIN	P53781 bos taurus
720	36	44.4	195	2	O02616_TRYCR	O02616 trypanosoma	793	36	44.4	326	2	Q5B3Y7_RAT	Q5b3y7 rattus norv
721	36	44.4	195	2	Q27791_TRYCR	Q27791 trypanosoma	794	36	44.4	330	1	A85B_MYCAV	Q06947 mycobacteri
722	36	44.4	195	2	Q71887_LEIMA	Q71887 leishmania	795	36	44.4	332	2	Q8KIQ1_PSEAE	Q8kiq1 pseudomonas
723	36	44.4	195	2	O4Q597_LEIMA	Q4q597 leishmania	796	36	44.4	334	2	Q742M3_MYCPA	Q742m3 mycobacteri
724	36	44.4	199	2	O87756_KLEPN	Q87756 klebsiella	797	36	44.4	335	2	O9S296_ARATH	Q9s296 arabidopsis
725	36	44.4	200	2	O4HGK1_CAMCO	O4hgk1 campylobact	798	36	44.4	339	2	O4WN85_ASPFU	Q4wn85 aspargillus
726	36	44.4	201	2	O5Y3J0_9NEOP	O5y3j0 pedicinus h	799	36	44.4	344	2	Q7YHM3_9NEOP	Q7yhm3 lepidopoecl
727	36	44.4	204	2	P91208_CAEEL	P91208 caenorhabdi	800	36	44.4	344	2	Q5PPG3_RAT	Q5ppg3 rattus norv
728	36	44.4	206	2	Q72D10_DESVH	Q72d10 desulfovibr	801	36	44.4	346	1	OXDA_RAT	Q35078 rattus norv
729	36	44.4	210	2	O8NCH4_HUMAN	O8nch4 homo sapien	802	36	44.4	347	2	Q9PT80_NORVI	Q9pt80 notophthalm
730	36	44.4	219	2	Q6D9S3_ERMCT	Q6d9s3 erwania car	803	36	44.4	349	2	Q6CRV3_KLULA	Q6crv3 kluyveromyc
731	36	44.4	220	2	O85NV1_9SAUR	O85nv1 pachydaetyl	804	36	44.4	359	2	O8JTH5_9CLOS	Q8jth5 little cher
732	36	44.4	221	2	O95719_LACAG	O95719 lacerta agi	805	36	44.4	360	2	O4NA72_9MCC	Q4na72 artibroacte
733	36	44.4	221	2	Q8M2D2_9SAUR	Q8m2d2 pachydaetyl	806	36	44.4	365	2	Q8WY01_HUMAN	Q8wy01 homo sapien
734	36	44.4	222	2	Q6SN77_9PRIM	Q6sn77 cercopithec	807	36	44.4	365	2	Q5YQK0_NOCFA	Q5yqk0 nocardia fa
735	36	44.4	223	2	Q7XJ22_ORYSA	Q7xj22 oryza sativ	808	36	44.4	370	1	TFFD1_ALCBU	P05404 alcaligenes
736	36	44.4	231	2	Q82MW7_STRAW	Q82mw7 streptomyc	809	36	44.4	371	2	Q8G9L1_RHOOP	Q8g9l1 rhodococcus
737	36	44.4	231	1	YC53L_SYNY3	P72583 synchocyst	810	36	44.4	373	2	Q8WXZ9_HUMAN	Q8wxz9 homo sapien
738	36	44.4	233	2	O6Z9K2_ORYSA	O6z9k2 oryza sativ	811	36	44.4	379	2	O585J6_9TRYP	Q585j6 trypanosoma
739	36	44.4	234	2	Q75CR4_ASHGO	Q75cr4 ashbya gos	812	36	44.4	381	2	Q9S4G4_9RODE	Q9s4g4 signodon le
740	36	44.4	235	2	Q6MPK0_BDEBA	Q6mpk0 bdellovibri	813	36	44.4	382	2	O99971_PORPU	Q99971 porphyra pu
741	36	44.4	237	2	Q6XNT0_9SAUR	Q6xnt0 phelsuma gu	814	36	44.4	385	2	Q6NG02_CORDI	Q6ng02 corynebacte
742	36	44.4	238	2	O9MP31_9HYME	O9mp31 dorymyrmex	815	36	44.4	389	2	Q7QVC2_GIALA	Q7qvc2 giardia lam
743	36	44.4	238	2	O6XNS9_9SAUR	O6xns9 phelsuma gu	816	36	44.4	390	2	O9S6K4_9RHIZ	Q9s6k4 agrobacteri
744	36	44.4	242	2	Q4T0B6_TETNG	Q4t0b6 tetraodon n	817	36	44.4	391	2	Q9TC99_NEPOL	Q9tc99 nephroselm
745	36	44.4	243	2	Q6D1D9_ERWCT	Q6d1d9 erwania car	818	36	44.4	393	2	O5PEM2_ANAMM	Q5pem2 anaplasma m
746	36	44.4	244	2	O5YJH8_9SAUR	O5yjh8 amphiglossu	819	36	44.4	394	2	Q88ZC8_LACPL	Q88zc8 lactobacill
747	36	44.4	249	2	Q4V8D7_RAT	Q4v8d7 rattus norv	820	36	44.4	398	2	Q7B5K6_LACPL	Q7b5k6 lactobacill
748	36	44.4	257	2	O5PX01_HCMV	O5px01 human cytom	821	36	44.4	400	2	O4SGC7_TETNG	Q4sgc7 tetraodon n
749	36	44.4	257	2	O5PX04_HCMV	O5px04 human cytom	822	36	44.4	402	2	O6BND0_DEBHA	Q6bnd0 debaryomyc
750	36	44.4	257	2	O5PX09_HCMV	O5px09 human cytom	823	36	44.4	406	2	Q5G534_PHOSU	Q5g534 phodopus eu
751	36	44.4	257	2	Q52N52_HCMV	Q52n52 human cytom	824	36	44.4	409	2	Q93CUB_SHIBO	Q93cub shigella bo
752	36	44.4	257	2	Q52N53_HCMV	Q52n53 human cytom	825	36	44.4	410	2	Q8NTI5_CORGL	Q8nti5 corynebacte
753	36	44.4	257	2	Q52N54_HCMV	Q52n54 human cytom	826	36	44.4	413	2	O541U0_RAT	Q541u0 rattus norv
754	36	44.4	257	2	Q52N55_HCMV	Q52n55 human cytom	827	36	44.4	417	2	O52CF4_MAGGR	Q52cf4 magnaporthe
755	36	44.4	258	2	Q89Q37_BRAJA	Q89q37 bradyrhizob	828	36	44.4	423	2	O528C4_MAGGR	Q528c4 magnaporthe
756	36	44.4	258	2	Q7TML5_MOUSE	Q7tml5 mus musculus	829	36	44.4	423	2	O5G4F9_9HIV1	Q5g4f9 human immun
757	36	44.4	258	2	O5PX08_HCMV	O5px08 human cytom	830	36	44.4	423	2	Q5G4C6_9HIV1	Q5g4c6 human immun
758	36	44.4	258	2	O5PX15_HCMV	O5px15 human cytom	831	36	44.4	428	1	NEUR3_BOVIN	Q97859 bos taurus
759	36	44.4	260	1	CCCP_DROME	Q76879 drosophila	832	36	44.4	428	2	Q4SGC9_TETNG	Q4sgc9 tetraodon n
760	36	44.4	260	2	Q5YML6_NOCFA	O5yml6 nocardia fa	833	36	44.4	434	2	O53EU6_HUMAN	Q53eu6 homo sapien
761	36	44.4	264	2	Q5YWL2_NOCFA	O5ywl2 nocardia fa	834	36	44.4	434	2	Q96NA3_HUMAN	Q96na3 homo sapien

835	36	44.4	439	2	Q9ELX0_9ALPH	Q9elx0 cercopithec	908	36	44.4	567	1	Y1354_MYCBO	P59372 mycobacteri
836	36	44.4	439	2	Q4TH55_TETNG	Q4th55 tetraodon n	909	36	44.4	567	2	Q4UDZ9_THEAN	Q4udz9 theileria a
837	36	44.4	442	2	Q52DK1_MAGGR	Q52dk1 magnaporth	910	36	44.4	568	2	Q8VCU1_MOUSE	Q8vcu1 mus musculu
838	36	44.4	444	2	Q60A44_METCA	Q60a44 methylococc	911	36	44.4	572	2	Q8YKM5_CERCA	Q8ykm5 ceratitlis c
839	36	44.4	445	1	HRH3_CAVPO	HRh3 cavia porce	912	36	44.4	575	2	Q8YK00_ANASP	Q8ykm0 anabaena sp
840	36	44.4	445	1	HRH3_HUMAN	Q9y5n1 homo sapien	913	36	44.4	575	2	Q8DXQ4_STRAS	Q8dxq4 streptococc
841	36	44.4	445	1	HRH3_MOUSE	P58406 mus musculu	914	36	44.4	575	2	Q8E3C3_STRAS	Q8e3c3 streptococc
842	36	44.4	445	1	HRH3_RAT	Q9qyn8 rattus norv	915	36	44.4	576	1	CATE_RHOCA	P37743 rhodobacter
843	36	44.4	445	2	Q548M6_HUMAN	Q548m6 homo sapien	916	36	44.4	580	2	Q5JCK2_AEDAL	Q5jck2 aedes albop
844	36	44.4	445	2	Q4QR17_HUMAN	Q4qr17 homo sapien	917	36	44.4	580	2	Q4ZST9_PSESY	Q4zst9 pseudomonas
845	36	44.4	445	2	Q865E1_MACMU	Q865e1 macaca mula	918	36	44.4	584	2	Q7SEB6_NEUCR	Q7seb6 neurospora
846	36	44.4	445	2	Q540P3_MOUSE	Q540p3 mus musculu	919	36	44.4	586	2	Q68TF0_DEBHA	Q68tf0 debaryomyce
847	36	44.4	445	2	Q5G535_PHOSU	Q5g535 phodopus su	920	36	44.4	588	2	Q4LRB6_9BURK	Q4lrb6 burkholderi
848	36	44.4	446	2	Q5F795_NEIGI	Q5f795 neisseria g	921	36	44.4	588	2	Q8YJVI_ANASP	Q8yjvi anabaena sp
849	36	44.4	446	2	Q5JYK2_NEIME	Q5jyk2 neisseria m	922	36	44.4	592	1	THD1_ARATH	Q9z866 arabidopsis
850	36	44.4	447	2	Q51XEO_MAGGR	Q51xe0 nocardia fa	923	36	44.4	592	2	Q8GUG5_ARATH	Q8gug5 arabidopsis
851	36	44.4	449	2	Q8NN88_CORGL	Q8nn88 magnaporth	924	36	44.4	597	2	Q25512_MANSE	Q25512 manduca sex
852	36	44.4	449	2	Q51XEO_MAGGR	Q51xe0 corynebacte	925	36	44.4	600	2	Q8CE00_MOUSE	Q8ce00 mus musculu
853	36	44.4	454	2	Q65AT1_9SPHN	Q65at1 sphingomon	926	36	44.4	604	2	Q6CC14_YARLI	Q6cc14 theileria a
854	36	44.4	454	2	Q67RK9_SYMTH	Q67rk9 symbiobacte	927	36	44.4	606	2	Q4WQ64_ASPFU	Q4wq64 aspergillus
855	36	44.4	455	1	GUDP_BACSU	P42237 bacillus su	928	36	44.4	613	2	Q9NJ10_TRINI	Q9nj10 rhizobium l
856	36	44.4	456	2	Q4KCE9_PSEF5	Q4kce9 pseudomonas	929	36	44.4	615	2	Q93L36_RHILV	Q93l36 mus musculu
857	36	44.4	457	2	Q4JXV1_CORJK	Q4jxv1 corynebacte	930	36	44.4	623	2	Q8BY18_MOUSE	Q8by18 mus musculu
858	36	44.4	459	2	Q9C772_ARATH	Q9c772 arabidopsis	931	36	44.4	627	2	Q25272_LEIDO	Q25272 leishmania
859	36	44.4	460	2	Q51WM3_MAGGR	Q51wm3 magnaporth	932	36	44.4	627	2	Q25294_LEIIN	Q25294 anopheles g
860	36	44.4	461	2	Q5KN29_CRYNE	Q5kn29 cryptococcu	933	36	44.4	631	2	Q7Q126_ANOGA	Q7q126 leishmania
861	36	44.4	467	2	Q7SAC8_NEUCR	Q7sac8 neurospora	935	36	44.4	631	2	Q9N6S9_LEIMA	Q9n6s9 leishmania
862	36	44.4	473	2	Q6ZS72_HUMAN	Q6zs72 homo sapien	936	36	44.4	634	2	Q7Q4C7_ANOGA	Q7q4c7 anopheles g
863	36	44.4	473	2	Q7UQ29_RHOBA	Q7uq29 rhodopirell	937	36	44.4	645	1	OPGH_XANCP	Q8p532 xanthomonas
864	36	44.4	474	2	Q972Y5_SULTO	Q972y5 sulfolobus	938	36	44.4	645	2	Q4UYZ8_XANCP	Q4uyz8 xanthomonas
865	36	44.4	474	2	Q51QNA_MAGGR	Q51qna4 magnaporth	939	36	44.4	645	2	Q5ZMQ9_CHICK	Q5zmq9 gallus gall
866	36	44.4	474	2	Q4I4F1_GIBZE	Q4i4f1 gibberella	940	36	44.4	649	2	Q75G80_ORYSA	Q75g80 oryza sativ
867	36	44.4	474	2	Q7QX87_GIALA	Q7qxe7 giardia lam	941	36	44.4	650	2	Q5YQ8_CRYNE	Q5yq8 cryptococcu
868	36	44.4	476	2	Q7QE72_ANOGA	Q7qe72 anopheles g	942	36	44.4	650	2	Q5YQ8_CRYNE	Q5yq8 cryptococcu
869	36	44.4	477	2	Q9RU05_DEIRA	Q9rug5 deinococcus	943	36	44.4	650	2	Q5YQ8_CRYNE	Q5yq8 cryptococcu
870	36	44.4	478	2	Q6ZHP8_ORYSA	Q6zhp8 oryza sativ	944	36	44.4	673	1	NAL10_MOUSE	Q8ccn1 mus musculu
871	36	44.4	484	2	Q83R75_SHIFL	Q83r75 shigella fl	945	36	44.4	673	1	UL06_HRV11	Q10130 human herpe
872	36	44.4	488	2	Q50E82_MOUSE	Q50e82 mus musculu	946	36	44.4	676	1	UL06_HRV11	Q89429 human herpe
873	36	44.4	491	2	Q8N115_HUMAN	Q8n115 homo sapien	947	36	44.4	678	1	UL06_HRV11	Q89429 human herpe
874	36	44.4	494	2	Q7XH54_ORYSA	Q7xh54 oryza sativ	948	36	44.4	688	2	Q802E3_MOUSE	Q802e3 mus musculu
875	36	44.4	494	2	Q8W5K4_ORYSA	Q8w5k4 oryza sativ	949	36	44.4	688	2	Q6PMU2_LEIIN	Q6pmu2 leishmania
876	36	44.4	498	2	Q9SZ45_ARATH	Q9sz45 arabidopsis	950	36	44.4	704	2	Q9UAC2_LEIDO	Q9uac2 leishmania
877	36	44.4	502	2	Q8BVF9_MOUSE	Q8bvf9 mus musculu	951	36	44.4	704	2	Q4QHH9_LEIMA	Q4qhh9 leishmania
878	36	44.4	503	2	Q9V7R7_DROME	Q9v7r7 drosophila	952	36	44.4	709	2	Q53QB3_ORYSA	Q53qb3 oryza sativ
879	36	44.4	505	2	Q8NA69_HUMAN	Q8na69 homo sapien	953	36	44.4	711	2	Q4QHH7_LEIMA	Q4qhh7 leishmania
880	36	44.4	513	2	Q93SD6_ECOLI	Q93sd6 escherichia	954	36	44.4	713	2	Q8MT14_LEITA	Q8mt14 leishmania
881	36	44.4	513	2	Q4FBH0_ECOLI	Q4fbh0 butyrivibri	955	36	44.4	720	2	Q9UAB8_LEIDO	Q9uab8 leishmania
882	36	44.4	517	1	XYLB_BUTFI	P45982 butyrivibri	956	36	44.4	722	2	Q6FN32_CANGA	Q6fn32 candida gla
883	36	44.4	519	2	Q9N3G2_CAEEL	Q9n3g2 caenorhabdi	957	36	44.4	724	2	Q7XH77_ORYSA	Q7xh77 oryza sativ
884	36	44.4	520	2	Q7QD18_ANOGA	Q7qdi8 anopheles g	958	36	44.4	725	2	Q8W5M3_ORYSA	Q8w5m3 oryza sativ
885	36	44.4	521	2	Q8H925_ORYSA	Q8h925 oryza sativ	959	36	44.4	725	2	Q8W5M3_ORYSA	Q8w5m3 oryza sativ
886	36	44.4	524	2	Q9KS18_VIBCH	Q9ks18 vibrio chol	960	36	44.4	728	2	Q51J30_MAGGR	Q51j30 magnaporth
887	36	44.4	524	2	Q91XD5_MOUSE	Q91xd5 mus musculu	961	36	44.4	732	2	Q4QHH8_LEIMA	Q4qhh8 leishmania
888	36	44.4	524	2	Q5ZLX4_CHICK	Q5zlx4 gallus gall	962	36	44.4	732	2	Q4QHH8_LEIMA	Q4qhh8 leishmania
889	36	44.4	527	2	Q51SD7_MAGGR	Q51sd7 magnaporth	963	36	44.4	749	2	Q6P943_BRARE	Q6p943 brachydanio
890	36	44.4	527	2	Q4M0E4_9BURK	Q4m0e4 burkholderi	964	36	44.4	766	2	Q8C4J2_MOUSE	Q8c4j2 mus musculu
891	36	44.4	528	2	Q4UVT6_XANCP	Q4uvt6 xanthomonas	965	36	44.4	771	2	Q4V782_MOUSE	Q4v782 mus musculu
892	36	44.4	528	2	Q8P8A1_XANCP	Q8p8a1 xanthomonas	966	36	44.4	775	2	Q7F9A6_ORYSA	Q7f9a6 oryza sativ
893	36	44.4	528	2	Q8PJQ6_XANAC	Q8pjq6 xanthomonas	967	36	44.4	776	2	Q9SU76_ARATH	Q9su76 arabidopsis
894	36	44.4	530	2	Q526K5_MAGGR	Q526k5 magnaporth	968	36	44.4	779	2	Q728Y0_DESVH	Q728y0 deesulfovibr
895	36	44.4	530	2	Q5GZ38_XANOR	Q5gz38 xanthomonas	969	36	44.4	786	2	Q75LD4_ORYSA	Q75ld4 oryza sativ
896	36	44.4	535	2	Q5LQW7_MAGGR	Q5lqm7 magnaporth	970	36	44.4	786	2	Q7XKS1_ORYSA	Q7xks1 oryza sativ
897	36	44.4	537	2	Q8XS18_RALSO	Q8xs18 raistonia s	971	36	44.4	789	2	Q75DB5_ASHGO	Q75db5 ashbya gos
898	36	44.4	538	2	Q92U07_RHIME	Q92u07 rhizobium m	972	36	44.4	816	2	Q4NTY9_9DELT	Q4nty9 anaeromyxob
899	36	44.4	541	2	Q9KZX5_STRCO	Q9kzx5 streptomyce	973	36	44.4	838	2	Q8K250_MOUSE	Q8k250 mus musculu
900	36	44.4	545	2	Q41FT6_GIBZE	Q41ft6 gibberella	974	36	44.4	845	2	Q8L4Q1_ORYSA	Q8l4q1 oryza sativ
901	36	44.4	551	2	Q7YWM0_HAEDU	Q7ymw0 haemophilus	975	36	44.4	852	2	Q6PB55_MOUSE	Q6pb55 mus musculu
902	36	44.4	557	2	Q51QW7_MAGGR	Q51qw7 magnaporth	976	36	44.4	857	2	Q4NN47_9DELT	Q4nn47 anaeromyxob
903	36	44.4	558	1	Y1145_VIBCH	Y1145 vibrio chol	977	36	44.4	885	1	ATG9_PICPA	Q876n4 pichia past
904	36	44.4	560	1	Y2420_PHOPE	Y2420 drosophila	978	36	44.4	922	2	Q8NAZ1_HUMAN	Q8nazi homo sapien
905	36	44.4	560	1	Y2420_PHOPE	Y2420 drosophila	979	36	44.4	923	2	Q4P3M5_5TMA	Q4p3m5 ustilago ma
906	36	44.4	562	2	Q4Q3B8_LEIMA	Q4q3b8 leishmania	980	36	44.4	927	2	Q4YZ18_PLABE	Q4yz18 plasmodium
907	36	44.4	567	1	Y1320_MYCTU	Q10633 mycobacteri							

981 36 44.4 953 2 05AS20_EMENI
982 36 44.4 953 2 08IYN5_HUMAN
983 36 44.4 967 2 05CEN2_CRYHO
984 36 44.4 1014 2 09MAR9_ARATH
985 36 44.4 1015 2 07XKG6_ORYSA
986 36 44.4 1068 2 06BJ02_DEBHA
987 36 44.4 1086 2 07SK50_ORYSA
988 36 44.4 1092 2 06PV87_ORYSA
989 36 44.4 1131 2 05LFG9_BACFN
990 36 44.4 1131 2 064W93_BACFR
991 36 44.4 1134 2 07X7G6_ORYSA
992 36 44.4 1142 2 04FY92_LEIMA
993 36 44.4 1158 2 09LNS1_ARATH
994 36 44.4 1176 2 07Y155_ORYSA
995 36 44.4 1184 2 055VR2_CRYNE
996 36 44.4 1184 2 05KN33_CRYNE
997 36 44.4 1190 2 09AYJ9_ORYSA
998 36 44.4 1191 2 09VUY8_DROME
999 36 44.4 1199 2 05LS54_MAGGR
1000 36 44.4 1204 2 0615B6_ORYSA

ALIGNMENTS

RESULT 1

Q6GNG6_XENLA PRELIMINARY; PRT; 508 AA.
AC Q6GNG6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC82822 protein.
GN Name=MGC82822;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC073546; AAH73546.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhGEF.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhGEF; 1.
DR PROSITE; PS0010; DH 2; 1.
DR PROSITE; PS0003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor.
SQ SEQUENCE 508 AA; 57858 MW; E8019B05639E93FE CRC64;
Query Match 61.7%; Score 50; DB 2; Length 508;
Best Local Similarity 46.2%; Pred. No. 7.9;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 2 ELVGSIMSWLTC 14
|::|::|::|::|
Db 214 EVIGPTMLEWLP 226
RESULT 2
Q4V789_XENTR PRELIMINARY; PRT; 526 AA.
ID Q4V789;
AC Q4V789;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC098078; AAH98078.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000219; RhGEF.

```
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RHOGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RHOGEF; 1.
DR PROSITE; PSS0010; DH_2; 1.
DR PROSITE; PSS0003; PH DOMAIN; 1.
KW Guanine-nucleotide releasing factor; Hypothetical protein.
SQ SEQUENCE 526 AA; 59250 MW; 1E6A5ADAF26C109F CRC64;

Query Match 61.78; Score 50; DB 2; Length 526;
Best Local Similarity 46.24; Pred. No. 8.2;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELVGPSLMSWLTC 14
DB 232 EVIGPTMLEWLP 244

RESULT 3
RNS11 HUMAN
ID RNS11_HUMAN STANDARD; PRT; 199 AA.
AC QRTAAL;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable ribonuclease 11 precursor (EC 3.1.27.-) (RNase 11).
GN Probable RNASE11, Synonyms=C14orf6; ORFNames=UNQ5832/PRO19669;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Placenta;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Testis;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togawa S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imobe N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
```

```

[3]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
Vandlen R.L., Watanabe C., Wiand D., Woods K., Xie M.-H.,
Yanura D.G., Yi S., Yu J., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).

[4]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: Belongs to the pancreatic ribonuclease family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX161458; CAD61920.1; -; mRNA.
CC EMBL; AK098687; BAC05378.1; -; mRNA.
CC EMBL; AY358794; AA089154.1; -; mRNA.
CC EMBL; BC025410; AAH25410.1; -; mRNA.
CC Ensembl; ENSG00000173464; Homo sapiens.
CC HGNC; HGNC:19269; RNASE11.
CC InterPro; IPR001427; RNaseA.
CC ProDom; PD000535; RNaseA; 1.
CC Endonuclease; Hydrolase; Nuclease; Signal.
FT SIGNAL 1 16 Potential.
FT CHAIN 17 199 Probable ribonuclease 11.
FT ACT SITE 82 82 Proton acceptor (By similarity).
FT SITE 115 119 Substrate binding (By similarity).
FT DISULFID 98 158 By similarity.
FT DISULFID 114 169 By similarity.
SQ SEQUENCE 199 AA; 22427 MW; BA9DBA907292F1F0 CRC64;

Query Match 60.5%; Score 49; DB 1; Length 199;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVGPSLMSWLTC 14
DB 184 LTGHSLSWLVC 195
```



```
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Aobayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Faragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003634; AAP53154.1; -; Genomic_DNA.
DR EMBL; AF273478; AAK58692.1; -; mRNA.
DR EMBL; AY051907; AAK93331.1; -; mRNA.
DR DR Ensembl; CG12317; Drosophila melanogaster.
DR FlyBase; FBgn028425; CGI12317.
DR FlyBase; FBgn028425; JHI-21.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
DR GO; GO:0006865; P:amino acid transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002293; AA/rel_permeasel.
DR InterPro; IPR004841; Permease_region.
DR Pfam; PF00324; AA_permease; 1.
DR Transmembrane; Transport.
KW SEQUENCE 500 AA; 54820 MW; 48BB44291114FEB8 CRC64;
SQ
Query Match 60.5%; Score 49; DB 2; Length 500;
Best Local Similarity 69.2%; Pred. No. 12; Mismatches 3; Indels 0; Gaps 0;
Matches 9; Conservative 1;

QY 2 ELVGPGLMSWLT 14
DB 71 ESVGSLLIWLT 83

RESULT 6
Q7XLV1_ORYSA
ID Q7XLV1_ORYSA PRELIMINARY; PRT; 1510 AA.
AC Q7XLV1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE OSJNBa0042115.9 protein.
GN Name=OSJNBa0042115.9;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RP MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL731641; CAE04887.3; -; Genomic DNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 1510 AA; 171089 MW; A9E138724D4ACD39 CRC64;
Query Match 60.5%; Score 49; DB 2; Length 1510;
Best Local Similarity 63.8%; Pred. No. 35; Mismatches 3; Indels 0; Gaps 0;
Matches 7; Conservative 1;

QY 2 ELVGPGLMSWLT 14
DB 71 ESVGSLLIWLT 83

RESULT 7
Q7XFO6_ORYSA
ID Q7XFO6_ORYSA PRELIMINARY; PRT; 167 AA.
AC Q7XFO6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSFNAMES=OSJNBb0008A05.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017076; AAP53011.1; -; Genomic_DNA.
DR Gramene; Q7XFO6; -.
KW Hypothetical protein.
SQ SEQUENCE 167 AA; 18151 MW; A217872B29BFBFB CRC64;
Query Match 58.0%; Score 47; DB 2; Length 167;
Best Local Similarity 61.5%; Pred. No. 8.6; Mismatches 4; Indels 0; Gaps 0;
Matches 8; Conservative 1;

QY 2 ELVGPGLMSWLT 14
DB 103 EAAEFSMASWLT 115

RESULT 8
Q8W5F2_ORYSA
ID Q8W5F2_ORYSA PRELIMINARY; PRT; 167 AA.
AC Q8W5F2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNBb0008A05.14.
GN Name=OSJNBb0008A05.14;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Ganeberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsetrin T., Riggs F., Heiao J., Zismann V., Blunt S., Pai G.,
RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091749; AAL31079.1; -; Genomic_DNA.
DR Gramene; Q8W5F2; -.
KW Hypothetical protein.
```

```
SQ SEQUENCE 167 AA; 18151 MW; A217872B29BFBFB CRC64;
Query Match 58.0%; Score 47; DB 2; Length 167;
Best Local Similarity 61.5%; Pred. No. 8.6;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ELVGPSSLMSWLT 14
Db 103 EAAEPSSMASWLT 115

RESULT 9
Q9HE61_NEUCR
ID Q9HE61_NEUCR PRELIMINARY; PRT; 745 AA.
AC Q9HE61
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Related to MKT1 protein.
GN Name=B2F7.010;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL451013; CAC18153.1; -; Genomic DNA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR006084; XPGC Rad.
DR PRINTS; PR00853; XPGRADSUPER.
DR SEQUENCE 745 AA; 84185 MW; 45F372292C354970 CRC64;

Query Match 59.0%; Score 47; DB 2; Length 745;
Best Local Similarity 63.6%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LVGPSSLMSWLT 13
Db 338 LIGPRLLSWIT 348

RESULT 10
O54512_YEREN
ID O54512_YEREN PRELIMINARY; PRT; 365 AA.
AC O54512
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Itp3 protein.
GN Name=irp3;
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=WA-314;
RX MEDLINE=98117033; PubMed=9457855;
RA Pelludat C., Rakin A., Jacobi C., Schubert S., Heesemann J.;
RT "The yersiniabactin biosynthetic gene cluster of Yersinia
enterocolitica: organization and siderophore-dependent regulation.";
RL J. Bacteriol. 180:538-546 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=WA-314;
```

```
RA Cosima P.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y12527; CAA73128.1; -; Genomic DNA.
DR PIR; T30343; T30343.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR Pfam; PF01408; GFO IDH MocA; 1.
DR TIGRFAMs; TIGR01761; thiaz-red; 1.
SQ SEQUENCE 365 AA; 40705 MW; E0CBB6EC56F3A7D1 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWL 12
Db 290 CETVGPEGVSWL 301

RESULT 11
Q9Z3C6_YERPE
ID Q9Z3C6_YERPE PRELIMINARY; PRT; 365 AA.
AC Q9Z3C6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE YbtU protein.
GN Name=ybtU;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=6/69;
RA Buchrieser C., Rusniok C., Couve E., Frangeul L., Billault A.,
RA Kunst F., Carniel E., Glaser P.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM6+;
RX MEDLINE=99035519; PubMed=9818149; DOI=10.1016/S1074-5521(98)90115-6;
RA Gehring A.M., DeMoll E., Fetherston J.D., Mori I., Mayhew G.F.,
RA Blatner F.R., Walsh C.T., Perry R.D.;
RT "Iron acquisition in plague: modular logic in enzymatic biogenesis of
Yersiniabactin by Yersinia pestis.";
RL Chem. Biol. 5:573-586 (1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=10231486;
RA Fetherston J.D., Bertolino V.J., Perry R.D.;
RT "YbtP and YbtQ: two ABC transporters required for iron uptake in
Yersinia pestis.";
RL Mol. Microbiol. 32:289-299 (1999).
DR EMBL; AL031866; CAA21392.1; -; Genomic DNA.
DR EMBL; AF091251; AAC69589.1; -; Genomic DNA.
DR PIR; T17441; T17441.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000683; GFO/IDH/MocA_N.
DR Pfam; PF01408; GFO IDH MocA; 1.
DR TIGRFAMs; TIGR01761; thiaz-red; 1.
SQ SEQUENCE 365 AA; 40782 MW; 3DA7E8320682B90E CRC64;

Query Match 56.8%; Score 46; DB 2; Length 365;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CELVGPSLMSWL 12
Db 290 CETVGPEGVSWL 301
```



```

DR InterPro; IPR010091; Thiaz_red.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
DR TIGRFAMs; TIGR01761; thiaz-red; 1.
SQ SEQUENCE 386 AA; 43080 MW; 35AB88FBBC72C827 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 386;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPGLMSWL 12
   ||||| :|||
Db 311 CETVGPGLMSWL 322

RESULT 14
O55WU3_CRYNE PRELIMINARY; PRT; 452 AA.
ID Q55WU3_CRYNE PRELIMINARY;
AC Q55WU3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBC0470;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OC NCBI_TaxID=283643;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=B-3501A;
RA Fung B., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
KW EMBL; AA050100013; EAL21906.1; -; Genomic_DNA.
DR Hypothetical protein.
SQ SEQUENCE 452 AA; 50503 MW; F3D6FB1D43FDF901 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 452;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CELVGPGLMSWL 12
   ||||| :|||
Db 320 CDLLGPGIVRWL 331

RESULT 15
Q5KJF5_CRYNE PRELIMINARY; PRT; 452 AA.
ID Q5KJF5_CRYNE PRELIMINARY;
AC Q5KJF5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Expressed protein.
GN ORFNames=CN06720;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OC NCBI_TaxID=214684;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RL EMBL; A5013844; AM85959.1; -; Genomic_DNA.
DR InterPro; IPR006683; GFO_IDH/MoCA_N.

RESULT 12
Q82F25_YERPE PRELIMINARY; PRT; 366 AA.
ID Q82F25_YERPE PRELIMINARY;
AC Q82F25;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Yersinia bioactive protein YbtU.
GN Name=YbtU; Synonyms=irp3; OrderedLocNames=YPO1909;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OC NCBI_TaxID=632;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL; AJ414150; CAC90725.1; -; Genomic_DNA.
DR PIR; A10232; A10232.
DR InterPro; IPR000583; GFO_IDH/MoCA_N.
DR InterPro; IPR010091; Thiaz_red.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
DR TIGRFAMs; TIGR01761; thiaz-red; 1.
KW Complete proteome.
SQ SEQUENCE 366 AA; 40912 MW; 400EE7E6097507E9 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 366;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CELVGPGLMSWL 12
   ||||| :|||
Db 291 CETVGPGLMSWL 302

RESULT 13
Q8D0C4_YERPE PRELIMINARY; PRT; 386 AA.
ID Q8D0C4_YERPE PRELIMINARY;
AC Q8D0C4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Thiazoliny1-S-HMWp1 reductase.
GN Name=ybtU; OrderedLocNames=y2401;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OC NCBI_TaxID=632;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=KIMS / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RL EMBL; A5013844; AM85959.1; -; Genomic_DNA.
DR InterPro; IPR006683; GFO_IDH/MoCA_N.
```

RA	Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D., Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E., Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J., D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J., Huang J.C., Jianbon G., Jones S.J.M., Koo H.L., Krzywinski M.I., Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.B., Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L., Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A., Suh B.B., Tenney A., Uterback T.R., Wickes B.L., Wortman J.R., Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W., Praser C.M., Hyman R.W.
RA	"the genome of the basidiomycetous yeast and human pathogen
RT	Cryptococcus neoformans."
RL	Science 307:1321-1324(2005).
RM	EMBL; AE017343; AA42646.1; -; Genomic_DNA.
RK	Complete proteome.
SQ	SEQUENCE 452 AA; 50504 MW; F3D6PB1D43PDPF901 CRC64;
Query Match	56.8%; Score 46; DB 2; Length 452;
Best Local Similarity	50.0%; Pred.No. 34;
Matches	6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy	1 CELVGPSLSMWL 12
Dd	320 CDLLGPGIVRWL 331
RESULT 16	
Q8VXB0_ORYSA	
ID	QBVBX0_ORYSA PRELIMINARY; PRT; 716 AA.
AC	QBVBX0;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DD	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Putative potassium transporter (fragment).
GN	Names=HAK33;
OS	Oryza sativa (Rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Erbartoideae; Oryzeae; Oryza.
OX	NCBI_TaxID=4530;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Banuelos M.A., Garciasblas B., Cubero B., Rodriguez-Navarro A.;
RT	"Inventory and functional characterization of the HAK potassium
RT	transporters of rice.";
RL	Plant Physiol. 130:784-795(2002).
RM	EMBL; AJ427982; CAD21003.1; -; Genomic_DNA.
DR	Gramene; QBVBX0; -.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0015079; F:potassium ion transporter activity; IEA.
DR	GO; GO:0006813; P:potassium ion transport; IEA.
DR	InterPro; IPR002355; Cu_ox_copper_BS.
DR	IPfam; PF02705; K+transpor.
DR	TIGRFAMs; TIGR00794; kup; 1.
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
FT	NON_TER 1
FT	NON_TER 716
SQ	SEQUENCE 716 AA; 78412 MW; A0391D3A41EEC2B5 CRC64;
Query Match	56.8%; Score 46; DB 2; Length 716;
Best Local Similarity	66.7%; Pred.No. 54;
Matches	8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy	3 LVGPISLSWLTC 14
Dd	202 LFAPILISWLTC 213
RESULT 17	
Q652J4_ORYSA	
ID	Q652J4_ORYSA PRELIMINARY; PRT; 778 AA.

RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysselis M., Karlsson E.,
RA Kells C., Kieu A., Kienner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Menes L.,
RA Mesirov J., Mihalov A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osmann S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Rogov P.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Sharpe T.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Smith C., Sougnez C.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Stavropoulos S.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Teamlia T., Tsomo N., Vallée D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR ENBL; AACP01000139; EAX85008.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1215 AA; 133959 MW; 7CD44C134F817739 CRC64;
Query Match 56.8%; Score 46; DB 2; Length 1215;
Best Local Similarity 63.6%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 4 VGPSLMSWLTC 14
Db 427 VGPKFQDWLTC 437
RESULT 19
Q6A7C7 PROAC PRELIMINARY; PRT; 305 AA.
AC Q6A7C7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Conserved phage-related protein.
GN OrderedLocusNames=PPA1599;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hostler F., Liesegang H., Wierzer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT83338.1; -; Genomic_DNA.
DR InterPro; IPR011664; Abi_2.
DR Pfam; PF07751; Abi_2; 1.
KW Complete proteome.
SQ SEQUENCE 305 AA; 35476 MW; 20CCE48EC9565472 CRC64;
Query Match 55.6%; Score 45; DB 2; Length 305;

Best Local Similarity 66.7%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 CELVGPQLMSWL 12
Db 197 CGLTGQPLESWL 208
RESULT 20
Q6NEV8 CORDI PRELIMINARY; PRT; 328 AA.
AC Q6NEV8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DIP2155;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Biotype gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cardeno-Tarraga A.-M., Estratiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holtroyd S., Jagels K., Moule S., Quail M.A.,
RA Hamlinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248360; CAE50683.1; -; Genomic_DNA.
DR InterPro; IPR011664; Abi_2.
DR Pfam; PF07751; Abi_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 328 AA; 37659 MW; 21B682B37C17D894 CRC64;
Query Match 55.6%; Score 45; DB 2; Length 328;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 CELVGPQLMSWL 12
Db 220 CGLTGQPLESWL 231
RESULT 21
Q8GAG2 ARTNI PRELIMINARY; PRT; 497 AA.
AC Q8GAG2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative amino acid permease.
OS Arthrobacter nicotinovorans.
OG Plasmid pAO1.
OC Bacteria; Actinobacteridae; Actinobacteriaceae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=29320;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95115562; PubMed=7815950;
RA Grether-Beck S., Igloi G.L., Pust S., Schiltz E., Decker K.,
RA Brandsch R.;
RT "Structural analysis and molybdenum-dependent expression of the pAO1-
RT encoded nicotine dehydrogenase genes of Arthrobacter nicotinovorans.";
RL Mol. Microbiol. 13:929-936(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96172783; PubMed=8588735; DOI=10.1007/s002030050246;

RA Menendez C., Igloi G., Henninger H., Brandsch R.;
 RT "A pAO1-encoded molybdopter in cofactor gene (moaA) of Arthrobacter
 RT nicotinovorans: characterization and site-directed mutagenesis of the
 RT encoded protein."; [3]
 RL Arch. Microbiol. 164:142-151 (1995).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98088982; PubMed=9428706;
 RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,
 RA Boettcher B., Brandsch R.;
 RT "Molybdate-uptake genes and molybdopter in-biosynthesis genes on a
 RT bacterial plasmid: characterization of MoaA as a filament-forming
 RT protein with adenosinetriphosphatase activity."; [4]
 RL Eur. J. Biochem. 250:524-531 (1997).
 RN [4]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99096870; PubMed=9878353; DOI=10.1006/jmbi.1998.2227;
 RA Schenk S., Hoelz A., Kraus B., Decker K.;
 RT "Gene structure and properties of enzymes of the plasmid-encoded
 RT nicotine catabolism of Arthrobacter nicotinovorans";
 RL J. Mol. Biol. 284:1323-1339 (1998).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97230479; PubMed=9073580; DOI=10.1006/plas.1996.1272;
 RA Menendez C., Igloi G.L., Brandsch R.;
 RT "TS1473, a putative insertion sequence identified in the plasmid pAO1
 RT from Arthrobacter nicotinovorans: isolation, characterisation and
 RT distribution among Arthrobacter species."; [6]
 RL Plasmid 37:35-41 (1997).
 RN [6]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21405725; PubMed=11514508;
 RX DOI=10.1128/JB.183.18.5262-5267.2001;
 RA Baitech D., Sandu C., Brandsch R., Igloi G.L.;
 RT "A gene cluster on pAO1 of Arthrobacter nicotinovorans involved in the
 RT degradation of the plant alkaloid nicotine: cloning, purification and
 RT characterization of 2,6-dihydroxypyridine 3-hydroxylase."; [7]
 RL J. Bacteriol. 183:5262-5267 (2001).
 RN [7]
 RN NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22505657; PubMed=12618462;
 RX DOI=10.1128/JB.185.6.1976-1986.2003;
 RA Igloi G.L., Brandsch R.;
 RT "Sequence of the 165-kilobase catabolic plasmid pAO1 from Arthrobacter
 RT nicotinovorans and identification of a pAO1-dependent nicotine uptake
 RT system."; [8]
 RL J. Bacteriol. 185:1976-1986 (2003).
 RL EMBL; AJ507836; CAD47969.1; -: Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR002293; AA/tel permease1.
 DR InterPro; IPR004841; Permease_region.
 DR Pfam; PF00324; AA_permease; 1.
 KW Plasmid; Transmembrane; Transport.
 SQ SEQUENCE 497 AA; 51904 MW; ED8804A8133B7B44 CRC64;
 Query Match 55.6%; Score 45; DB 2; Length 497;
 Best Local Similarity 63.6%; Pred. No. 56;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LVGPSLMSWLT 13
 Db 381 LIGPALLWLT 391
 RESULT 22
 ID ARHG3 MOUSE
 AC Q91X46; Q8CDM0; Q91V4; Q99K14; Q9DC31;
 DT 25-OCT-2004 (Rel. 45, Created)

DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Rho guanine nucleotide exchange factor 3.
 GN Name=Arhgef3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 2 AND 3).
 RP STRAIN=C57BL/6J; TISSUE=Lung, and Testis;
 RX MEDLINE=22354663; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelain A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang L., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai K., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."; [2]
 RL Nature 420:563-573 (2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RP STRAIN=Czech 11, and FVB/N; TISSUE=Mammary tumor, and Salivary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."; [3]
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RN FUNCTION, AND SUBCELLULAR LOCATION.
 RP PubMed=11839749; DOI=10.1074/jbc.M111108200;
 RA Schmidt A., Hall A.;
 RT "The Rho exchange factor Net1 is regulated by nuclear sequestration.";
 RL J. Biol. Chem. 277:14581-14588 (2002).
 CC -!- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for

```
CC RhoA and RhoB GTPases.
CC -|- SUBUNIT: Interacts with RHOA and RHOB.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q91X46-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q91X46-2; Sequence=VSP_011613;
CC Name=3;
CC IsoId=Q91X46-3; Sequence=VSP_011614;
CC -|- SIMILARITY: Contains 1 PH (DBL-homology) domain.
CC -|- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AK004600; BAB23401.1; -; mRNA.
CC EMBL; AK029872; BAC26652.1; -; mRNA.
CC EMBL; BC005517; AAH05517.1; -; mRNA.
CC EMBL; BC007153; AAH07153.1; ALT INIT; mRNA.
CC EMBL; BC012262; AAH12262.1; -; mRNA.
CC HSP; Q15811.1; K11
CC Ensembl; ENSMUSG00000021895; Mus musculus.
CC MGI; MGI:1918954; Arhgef3.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH type.
CC InterPro; IPR00219; RHOGEF.
CC Pfam; PF00169; PH; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RHOGEF; 1.
CC PROSITE; PS50010; DH 2; 1.
CC PROSITE; PS50003; PH DOMAIN; 1.
CC Alternative splicing; Guanine-nucleotide releasing factor.
CC DOMAIN 121 303
CC DOMAIN 290 448
CC VARSPLIC 1 31
CC MVAKDYPFLTVKRNCSLEAPLGSVAKDE -> MRSERP
CC MVWCCFFLRAQKRKQSQDEDAVSLCSLDIS (in
CC isoform 2)
CC /FTId=VSP_011613.
CC MVAKDYPFLTVKRNCSLEAPLGSVAKDE -> MFPSPK
CC ACNFRGRKRKQSQDEDAVSLCSLDIS (in isoform
CC 3).
CC /FTId=VSP_011614.
CC V -> I (in Ref. 2; AAH05517).
CC CONFLICT 363 363 V -> I (in Ref. 2; AAH05517).
CC SEQUENCE 524 AA; 59526 MW; 579FP96C1F0B0ADA CRC64;
Query Match 55.6%; Score 45; DB 1; Length 524;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 2 ELVGSPLMSWLTC 14
Db 194 EHVGPILVGLPC 206
RESULT 23
ARHG3 HUMAN STANDARD; PRT; 526 AA.
AC Q9NR81; Q9NRJ3; Q724U2; Q725T2; Q9H7T4;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rho guanine nucleotide exchange factor 3 (Exchange factor found in
DE platelets and leukemic and neuronal tissues) (XPLN).
GN Name=ARHG3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
```

```
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=20334277; PubMed=10873612; DOI=10.1006/bbr.2000.2925;
RA Thiesen S., Kuebart S., Ropers H.-H., Notthwang H.G.;
RT "Isolation of two novel human RHOGEFs, ARHGEP3 and ARHGEP4, in 3p13-21
RL Biochem. Biophys. Res. Commun. 273:364-369(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glasel S.,
RA Ansgar W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2), AND VARIANT VAL-335.
RX Guo J.H., Yu L.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1), AND VARIANT
RL VAL-335.
RC TISSUE=Brain, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 71-526 (ISOFORMS 1/2).
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu K., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
```

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohnori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [6]
RP FUNCTION AND TISSUE SPECIFICITY.
RX PubMed=1221096; DOI=10.1074/jbc.M207401200;
RA Arthur W.T., Ellertbroek S.M., Der C.J., Burridge K., Wennerberg K.;
RT "XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not
RT RhoC.";
RL J. Biol. Chem. 277:42964-42972(2002).
CC -I- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for
CC RhoA and RhoB GTPases.
CC -I- SUBUNIT: Interacts with RHOA and RHOB.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Cyping; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NR81-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NR81-2; Sequence=VSP_011612;
CC -I- TISSUE SPECIFICITY: Widely expressed. Highest levels are found in
CC adult brain and skeletal muscle. Lower levels are found in heart
CC and kidney.
CC -I- SIMILARITY: Contains 1 DH (DBL-homology) domain.
CC -I- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF249744; AAF79954.1; -; mRNA.
CC EMBL; AL136832; CAB66766.1; -; mRNA.
CC EMBL; AF433862; AAP97313.1; -; mRNA.
CC EMBL; BC054345; AAH54345.1; ALT_INIT; mRNA.
CC EMBL; BC068513; AAH68513.1; ALT_INIT; mRNA.
CC EMBL; AK024340; BAB14891.1; ALT_INIT; mRNA.
CC HSSP; Q15811; IKIL.
CC Ensembl; ENSG00000163947; Homo sapiens.
CC HGNC; HGNC:683; ARHGEF3.
CC GO; GO:0005089; F:Rho guanyl-nucleotide exchange factor activity; TAS.
CC GO; GO:0007266; P:Rho protein signal transduction; TAS.
CC InterPro; IPR001331; GDS_CDC24.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH_type.
CC InterPro; IPR000219; RhoGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00621; RhoGEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RhoGEF; 1.
CC SMART; PS00741; DH_1; FALSE_NEG.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC Alternative splicing; Guanine-nucleotide releasing factor;
KW Polymorphism.
KW DOMAIN 122 304 DH.
FT DOMAIN 291 449 PH.
FT VARSP LIC 1 32 MVAKDYFFYIVTVKANCSELPASGPAKDAAE -> MDSSST
FT ANMQCSGRGMEENKRPKQRQNNFMFPSPKAWNFRGRKR
FT KQSTQDEDAVSLGSLDIS (in isoform 2).
FT /FTId=VSP_011612.
FT K -> R (in dbSNP:3732507).
FT /FTId=VAR_021935.
FT VARIANT 13 13

FT VARIANT 335 335 L -> V (in dbSNP:3772219).
FT /FTId=VAR_021936.
FT CONFLICT 410 410 I -> T (in Ref. 3).
SQ SEQUENCE 526 AA; 59783 MW; 0FE8249AB52D1C8C CRC64;
Query Match 55.6%; Score 45; DB 1; Length 526;
Best Local Similarity 61.5%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 ELVGPSLMSWLTC 14
Db 195 EHVGPILVGLWLC 207

RESULT 24
ARHG3_MACFA
ID ARHG3_MACFA STANDARD; PRT; 526 AA.
AC Q9NOA8;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rho guanine nucleotide exchange factor 3.
GN Name=ARHGEF3; ORFNames=QcCE-16434;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheciae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Acts as guanine nucleotide exchange factor (GEF) for
CC RhoA and RhoB GTPases (By similarity).
CC -I- SUBUNIT: Interacts with RHOA and RHOB (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Contains 1 DH (DBL-homology) domain.
CC -I- SIMILARITY: Contains 1 PH domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB046022; BAB01604.1; -; mRNA.
CC HSSP; Q15811; IKIL.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011993; PH_type.
CC InterPro; IPR000219; RhoGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00621; RhoGEF; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00325; RhoGEF; 1.
CC PROSITE; PS50010; DH_2; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor.
FT DOMAIN 122 304 DH.
FT DOMAIN 291 449 PH.
SQ SEQUENCE 526 AA; 59764 MW; 0FE83450D52D1C8C CRC64;
Query Match 55.6%; Score 45; DB 1; Length 526;
Best Local Similarity 61.5%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 ELVGPSLMSWLTC 14
Db 195 EHVGPILVGLWLC 207

```
RESULT 25
QSR6F2_PONPY PRELIMINARY; PRT; 526 AA.
ID QSR6F2_PONPY PRELIMINARY; PRT; 526 AA.
AC QSR6F2;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFp459N027.
GN Name=DKFp459N027;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The German cDNA Consortium;
RA Ottenwelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RA Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; CR860538; CAH92664.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR00219; RhoGEF.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW GTPase activation; Guanine-nucleotide releasing factor;
KW Hypothetical protein.
SQ SEQUENCE 526 AA; 59811 MW; 01C12AAD8B041CA5 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 526;
Best Local Similarity 61.5%; Pred. No. 60;
Matches 8; Conservative 1; Mismatches 1; Indels 4; Gaps 0;

Qy 2 ELVGPILVMSWLTG 14
Db 195 EHVGPILVGLPC 207

RESULT 26
Q59F00_HUMAN PRELIMINARY; PRT; 537 AA.
ID Q59F00_HUMAN PRELIMINARY; PRT; 537 AA.
AC Q59F00;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Rho guanine nucleotide exchange factor 3 variant (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RA "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB209661; BAD92898.1; -; mRNA.
FT NON_TER 1
SQ SEQUENCE 537 AA; 61407 MW; 870802D1D27D65D3 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 537;
Best Local Similarity 61.5%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 4; Indels 4; Gaps 0;

Qy 2 ELVGPILVMSWLTG 14
Db 235 EHVGPILVGLPC 247

RESULT 27
Q4QQ00_HUMAN PRELIMINARY; PRT; 566 AA.
ID Q4QQ00_HUMAN PRELIMINARY; PRT; 566 AA.
AC Q4QQ00;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=PCR rescued clones;
RA MDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Dickinson M.C.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.B.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=PCR rescued clones;
RG NIH MGC Project;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC098122; AAH98122.1; -; mRNA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR00219; RhoGEF.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR PROSITE; PS50010; DH 2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 566 AA; 63745 MW; 43B89906644FB89D CRC64;

Query Match 55.6%; Score 45; DB 2; Length 566;
Best Local Similarity 61.5%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 4; Indels 4; Gaps 0;

Qy 2 ELVGPILVMSWLTG 14
Db 235 EHVGPILVGLPC 247

RESULT 28
Q4QQ15_HUMAN PRELIMINARY; PRT; 566 AA.
ID Q4QQ15_HUMAN PRELIMINARY; PRT; 566 AA.
AC Q4QQ15;
```

DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE-PCR rescued clones;
 RC MDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Datchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., Morley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley P.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC NTH MGC Project;
 RG TISSUE-PCR rescued clones;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; BC098272; AAH98272.1; PH.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RhoGEF.
 DR Pfam; PF00169; PH; 1.
 DR SMART; SM00621; RhoGEF; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RhoGEF; 1.
 DR PROSITE; PSS0010; DH 2; 1.
 DR PROSITE; PSS0003; PH DOMAIN; 1.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 566 AA; 63759 MW; D45197FD63AFAD72 CRC64;
 Query Match 55.6%; Score 45; DB 2; Length 566;
 Best Local Similarity 61.5%; Pred. No. 64;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ELVGPMSLWLC 14
 Db |||||:
 235 EHVGPILVGLWC 247
 RESULT 29
 Q52BN4_MAGGR PRELIMINARY; PRT; 734 AA.
 AC Q52BN4;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG08303.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha.

OX NCBI_TaxID=242507;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Alt-zahra M., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Ambruster J., Bachanteang P., Baldwin J., Barry A.,
 RA Bayul T., Blitshteyn B., Bloom T., Bye J., Boguslavskiy L.,
 RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawoe T., Degray N., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H., Gierre S.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
 RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltzen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvasselis M., Karlsson E.,
 RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
 RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
 RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
 RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
 RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Meneus L.,
 RA Mesirov J., Mihalov A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
 RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
 RA Norbu N., O'donnell P., Okawo O., O'leary S., Omotosho B.,
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
 RA Ratta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
 RA Rutman M., Schuppach R., Seaman C., Settippalli S., Sharpe T.,
 RA Spencer J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
 RA Stetson K., Stalker J., Stange-thomann N., Stavropoulos P.,
 RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
 RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
 RA Zimmer A., Zody M., Lander E.;
 RT "The genome sequence of Magnaporthe grisea";
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACU01000328; EAA57334.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 734 AA; 81983 MW; 316892F03309C9A5 CRC64;
 Query Match 55.6%; Score 45; DB 2; Length 734;
 Best Local Similarity 54.5%; Pred. No. 83;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LVGPMSLWLT 13
 Db |||||:
 339 LIGPRMLSWIT 349
 RESULT 30
 Q8VZ60_ANASP PRELIMINARY; PRT; 907 AA.
 ID Q8VZ60_ANASP

Mon May 15 11:35:41 2006

```

AC Q8YZ60;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Alr0627 protein.
GN OrderedLocusNames=alr0627;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; BA000019; BAB72595.1; -; Genomic_DNA.
DR PIR; AB1885; AB1885.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR004300; Glyco_hydro_57.
DR Pfam; PF03065; Glyco_hydro_57; 1.
KW Complete proteome.
SQ SEQUENCE 907 AA; 104134 MW; 5F0044217B8EB39 CRC64;

Query Match 55.6%; Score 45; DB 2; Length 907;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VGPSLMSWL 12
Db 106 IGPTLMSWL 114

RESULT 31
Q5TM27 MACMU
ID Q5TM27 MACMU PRELIMINARY; PRT; 481 AA.
AC Q5TM27;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript-3 (Fragment).
GN Name=BAT3;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheciidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kuleki J.K., Anzai T., Shiina T., Inoko H.;
RT "Rhesus Macaque Class I Dupliron Structures, Organization, and
RT Evolution Within the Alpha Block of the Major Histocompatibility
RT Complex.";
RL Mol. Biol. Evol. 21:2079-2091(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fukami-Kobayashi K., Shiina T., Anzai T., Sano K., Yamazaki M.,
RA Inoko H., Tateno Y.;
RT "Genomic evolution of MHC class I region in primates.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:9230-9234(2005).
DR EMBL; AB128049; BAD69718.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 481 AA; 51932 MW; 1CDDC0CB4688DF89 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 481;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 4 VGPSLMSWL 13
Db 255 VNPSLSWL 264

RESULT 32
Q8GA10 ARTNI
ID Q8GA10 ARTNI PRELIMINARY; PRT; 481 AA.
AC Q8GA10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative amino acid permease.
OS Arthrobacter nicotinovorans.
OG Arthrobacter nicotinovorans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=29320;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96172783; PubMed=8588735; DOI=10.1007/s002030050246;
RA Menendez C., Igloi G., Henninger H., Brandsch R.;
RT "A pAOI-encoded molybdopterin cofactor gene (moaA) of Arthrobacter
RT nicotinovorans: characterization and site-directed mutagenesis of the
RT encoded protein.";
RL Arch. Microbiol. 164:142-151(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9808982; PubMed=9428706;
RA Menendez C., Otto A., Igloi G., Nick P., Brandsch R., Schubach B.,
RA Boettcher B., Brandsch R.;
RT "Molybdate-uptake genes and molybdopterin-biosynthesis genes on a
RT bacterial plasmid: characterization of MoaA as a filament-forming
RT protein with adenosinetriphosphatase activity.";
RL Eur. J. Biochem. 250:524-531(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99096870; PubMed=9878353; DOI=10.1006/jmbi.1998.2227;
RA Schenk S., Hoelz A., Kraus B., Decker K.;
RT "Gene structure and properties of enzymes of the plasmid-encoded
RT nicotine catabolism of Arthrobacter nicotinovorans.";
RL J. Mol. Biol. 284:1323-1339(1998).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97230479; PubMed=9073580; DOI=10.1006/plas.1996.1272;
RA Menendez C., Igloi G.L., Brandsch R.;
RT "IS1473, a putative insertion sequence identified in the plasmid pAOI
RT from Arthrobacter nicotinovorans: isolation, characterisation and
RT distribution among Arthrobacter species.";
RL Plasmid 37:35-41(1997).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21405725; PubMed=11514508;
RA Baitsch D., Sandu C., Brandsch R., Igloi G.L.;
RT "A gene cluster on pAOI of Arthrobacter nicotinovorans involved in the
RT degradation of the plant alkaloid nicotine: cloning, purification and
RT characterization of 2,6-dihydroxypyridine 3-hydroxylase.";
RL J. Bacteriol. 183:5262-5267(2001).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22505657; PubMed=12618462;
RA Igloi G.L., Brandsch R.;
RT "Sequence of the 16S-kilobase catabolic plasmid pAOI from Arthrobacter

```

```

RT nicotinovorans and identification of a PAOI-dependent nicotine uptake
RT system."
RL J. Bacteriol. 195:1976-1986(2003).
DR EMBL: AJ507836; CAD47924.1; -; Genomic DNA.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016020; C: membrane; IEA.
DR GO: GO:0005279; F: amino acid-polyamine transporter activity; IEA.
DR GO: GO:0006865; F: amino acid transport; IEA.
DR GO: GO:0006810; F: transport; IEA.
DR InterPro: IPR002293; AA/permease1.
DR InterPro: IPR004841; Permease region.
DR Pfam: PF00324; AA_permease_1.
KW Plasmid; Transmembrane; Transport.
SQ SEQUENCE 481 AA; 49783 MW; 4EA9FB3BB8B76B64 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 481;
Best Local Similarity 72.7%; Pred. No. 81;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LVGPPLMSWLT 13
Db 367 LVGPPLMSWLT 377
||||:|||||

RESULT 33
ID Q504F6_BRARE PRELIMINARY; PRT; 506 AA.
AC Q504F6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein zgc:109904.
GN Names=zgc:109904;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL: BC095043; AAH95043.1; -; mRNA.
DR InterPro: IPR001849; PH.
DR InterPro: IPR011993; PH_type.
DR InterPro: IPR000219; RhogGFP.
DR InterPro: IPR012335; Thioredoxin_fold.

```

```

DR Pfam: PF00169; PH; 1.
DR Pfam: PF00621; RhogGFP; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00325; RhogGFP; 1.
DR PROSITE: PS00010; DH_2; 1.
DR PROSITE: PS00003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor; Hypothetical protein.
SQ SEQUENCE 506 AA; 57615 MW; 1F087D893CBB7EF CRC64;

Query Match 54.3%; Score 44; DB 2; Length 506;
Best Local Similarity 54.5%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 14
Db 210 VGPTMDWPC 220
||||:|||||

RESULT 34
ID Q8SNA3_MOUSE PRELIMINARY; PRT; 836 AA.
AC Q8SNA3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bat3 protein (Fragment).
GN Name=Bat3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC026647; AAH26647.1; -; mRNA.
DR MGI: MGI:1319439; Bat3.
DR GO: GO:0005515; F: protein binding; IPL.
FT NON_TER 1
SQ SEQUENCE 836 AA; 86873 MW; E2ED17337D724E71 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 836;
Best Local Similarity 80.0%; Pred. No. 144+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPSLMSWLT 13
Db 4 VGPSLMSWLT 13
||||:|||||

```

```
Db          610 VNPSLVSWLT 619

RESULT 35
Q5SQ40 HUMAN
ID Q5SQ40 HUMAN PRELIMINARY; PRT; 1029 AA.
AC Q5SQ40;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-013;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18505.1; -; Genomic DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1029 AA; 108672 MW; B694B35E126B3091 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1029;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy          4 VGPSLMSWLT 13
| | | | |
Db          900 VNPSLVSWLT 909

RESULT 36
Q5SQ42 HUMAN
ID Q5SQ42 HUMAN PRELIMINARY; PRT; 1077 AA.
AC Q5SQ42;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-019;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18503.1; -; Genomic DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1077 AA; 113455 MW; CB94D1C58A1E3D21 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1077;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy          4 VGPSLMSWLT 13
| | | | |
Db          900 VNPSLVSWLT 909

RESULT 37
Q5SQ43 HUMAN
ID Q5SQ43 HUMAN PRELIMINARY; PRT; 1078 AA.
AC Q5SQ43;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3.
GN Name=BAT3; ORFNames=DAQB-195H10.3-020;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CAI18502.1; -; Genomic DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1078 AA; 113909 MW; 4CBCB9F9FE02CF79 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1078;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy          4 VGPSLMSWLT 13
| | | | |
Db          900 VNPSLVSWLT 909

RESULT 38
Q6MG49 RAT
ID Q6MG49 RAT PRELIMINARY; PRT; 1096 AA.
AC Q6MG49;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HLA-B associated transcript 3, rat orthologue.
GN Name=Bat3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Brown Norway;
RC PubMed=15060004; DOI=10.1101/gr.1987704;
RA Hurt P., Walter L., Sudbrak R., Klages S., Mueller I., Shiina T.,
RA Inoko H., Lehrach H., Guenther E., Reinhardt R., Himmelbauer H.;
RT "The genomic sequence and comparative analysis of the rat major
RT histocompatibility complex.";
RL Genome Res. 14:631-639(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Brown Norway;
RA Boehm S., Borzym K., Thiel J., Sontag M., Hurt P., Himmelbauer H.,
RA Lang N., Lehrach S.,
RA Sudbrak R., Reinhardt R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX83045; CAE83997.1; -; Genomic DNA.
DR HSSP; Q862M4; IAAK.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
```

```
DR SMART; SM00213; UBG; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
SQ SEQUENCE 1096 AA; 114646 MW; E372D70E024DBF15 CRC64;

Query Match      54.3%; Score 44; DB 2; Length 1096;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VGPSLMSWLT 13
Db      919 VNPSLSWLT 928

RESULT 39
Q9WTN8_RAT PRELIMINARY; PRT; 1098 AA.
AC Q9WTN8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3.
GN Name=BAT3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99317047; PubMed=10390159; DOI=10.1089/10454999315222;
RA Ozaki T., Hanaoka E., Naka M., Nakagawara A., Sakiyama S.;
RT "Cloning and characterization of rat BAT3 cDNA."
RL DNA Cell Biol. 18:503-512(1999).
DR EMBL; AB018791; BAA76607.1; -. mRNA.
DR HSSP; Q9SHE7; 1BT0.
DR Ensembl; ENSRNOG0000000851; Rattus norvegicus.
DR RGD; 71064; Bat3.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBG; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
SQ SEQUENCE 1098 AA; 114931 MW; 5246PFE7D5CD8FB3 CRC64;

Query Match      54.3%; Score 44; DB 2; Length 1098;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VGPSLMSWLT 13
Db      921 VNPSLSWLT 930

RESULT 40
Q9HYL9_HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q9HYL9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein DKFP686L0653.
GN Name=DKFP686L0653;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Endometrium;
RG The German cDNA Consortium;
```

```
RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Pobo G., Han M., Wiemann S.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX647244; CAI46045.1; -. mRNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBG; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1126 AA; 118724 MW; 967ABE046DED3ABE CRC64;

Query Match      54.3%; Score 44; DB 2; Length 1126;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VGPSLMSWLT 13
Db      900 VNPSLSWLT 909

RESULT 41
Q96SA6_HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q96SA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3 protein.
GN Name=BAT3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shiina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; RA000025; BAB63390.1; -. Genomic_DNA.
DR HSSP; Q9SHE7; 1BT0.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBG; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
SQ SEQUENCE 1126 AA; 118702 MW; 8A67290B45B46ABA CRC64;

Query Match      54.3%; Score 44; DB 2; Length 1126;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      4 VGPSLMSWLT 13
Db      900 VNPSLSWLT 909

RESULT 42
Q9BCN4_HUMAN PRELIMINARY; PRT; 1126 AA.
AC Q9BCN4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE HLA-B associated transcript-3, isoform b.
GN Name=BAT3;
```

GN ORFNames=DAOB-195H10.3-001, XXbac-BCX270M2.3-001,
GN XXbac-BPC298P20.4-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Tracey A.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Griffiths C.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Wood J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003133; AA03133.1; -; Genomic DNA.
DR EMBL; AL662801; CA118315.1; -; Genomic DNA.
DR EMBL; AL662847; CA117658.1; -; Genomic DNA.
DR EMBL; AL670886; CA117784.1; -; Genomic DNA.
DR EMBL; AL805934; CA118501.1; -; Genomic DNA.
DR EMBL; AL670886; CA117658.1; JOINED; Genomic DNA.
DR EMBL; AL662847; CA117784.1; JOINED; Genomic DNA.
DR HSSP; Q9SH57; I8T0.
DR Ensembl; ENSG00000096155; Homo sapiens.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1126 AA; 118692 MW; 8A672908C8176ABA CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1126;

Best Local Similarity 80.0%; Pred. No. 1.9e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 8; Conservative 1; Indels 0; Gaps 0;

QY 4 VGPSLSWLT 13

|||||

900 VNPSLSWLT 909

Db

RESULT 43
BAT3_HUMAN
ID BAT3_HUMAN STANDARD; PRT; 1132 AA.
AC P46379; 1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Large proline-rich protein BAT3 (HLA-B-associated transcript 3) (G3).
GN Name=BAT3; Synonyms=G3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=i-cell;
RX MEDLINE=90192810; PubMed=2156268;
RA Banerji J., Sands J., Strominger J.L., Spies T.;
RT "A gene pair from the human major histocompatibility complex encodes
RT large proline-rich proteins with multiple repeated motifs and a single
RT ubiquitin-like domain";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
RN [2]
RP PHOSPHORYLATION SITES SER-964; SER-973 AND SER-1117.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
RA Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
CC -!- FUNCTION: Unknown.
CC -!- INTERACTION:
CC Q96WS7; -: NExp=1; IntAct=EBI-347552, EBI-372406;
CC Q9NWXI; RNF126; NExp=1; IntAct=EBI-347552, EBI-347629;
CC Q96S82; SBI32; NExp=1; IntAct=EBI-347552, EBI-348604;
CC Q43765; SGT4; NExp=1; IntAct=EBI-347552, EBI-347996;
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; M33519; AAA35587.1; -; mRNA.
DR EMBL; M33521; AAA35588.1; -; Genomic DNA.
DR EMBL; M33520; AAA35588.1; JOINED; Genomic DNA.
DR PIR; A35098; A35098.
DR HSSP; P02248; IUD7.
DR IntAct; P46379; -.
DR Ensembl; ENSG00000096155; Homo sapiens.
DR HGNC; HGNC:13919; BAT3.
DR H-InvDB; HIX0005722; -.
DR MIM; 142590; -.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR Phosphorylation; Repeat.
KW DOMAIN 17 77 Ubiquitin-like.
FT REPEAT 242 270 1.
FT REPEAT 415 443 2.
FT REPEAT 574 602 3.
FT REPEAT 608 636 4.
FT REGION 242 636 4 X 29 AA approximate repeats.
FT COMPIAS 202 207 Poly-Pro.
FT COMPIAS 657 670 Poly-Pro.
FT MOD_RES 964 964 Phosphoserine.
FT MOD_RES 973 973 Phosphoserine.
FT MOD_RES 1117 1117 Phosphoserine.
SQ SEQUENCE 1132 AA; 119504 MW; E28CABA78C38DD18 CRC64;

Query Match 54.3%; Score 44; DB 1; Length 1132;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VGPFLMSWLT 13
 DB 906 VNPSLVSWLT 915

RESULT 44
 Q5STC1 HUMAN
 ID Q5STC1_HUMAN PRELIMINARY; PRT; 1132 AA.
 AC Q5STC1
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE HLA-B associated transcript 3.
 GN Name=BAT3;
 GN ORFNames=DAQB-195H10.3-002, XXBac-BCX270M2.3-002,
 GN XXBac-BPG296P20.4-002;
 GN OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Tracey A.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Griffiths C.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Wood J.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL662847; CA117859.1; -; Genomic DNA.
 DR EMBL; AL670886; CA117785.1; -; Genomic DNA.
 DR EMBL; AL805934; CA118504.1; -; Genomic DNA.
 DR EMBL; AL662801; CA118314.1; -; Genomic DNA.
 DR EMBL; AL670886; CA117659.1; JOINED; Genomic DNA.
 DR EMBL; AL662847; CA117785.1; JOINED; Genomic DNA.
 DR Ensembl; ENSG00000096155; Homo sapiens.
 DR GO; GO:0006464; P:protein modification; IEA.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS00299; UBIQUITIN_1; 1.
 DR PROSITE; PS00053; UBIQUITIN_2; 1.
 SQ SEQUENCE 1132 AA; 119409 MW; 625B5F86321367ED CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1132;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 VGPFLMSWLT 13
 DB 906 VNPSLVSWLT 915

RESULT 45
 Q92LR2 MOUSE
 ID Q92LR2_MOUSE PRELIMINARY; PRT; 1154 AA.
 AC Q92LR2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE BAT3.
 GN Name=Bat3;
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=129;
 RX PubMed=14656967; DOI=10.1101/gr.1736803;
 RA Xie T., Rowen L., Aguado B., Ahearn M.E., Madan A., Qin S.,
 RA Campbell R.D., Hood L.;
 RT "Analysis of the gene-dense major histocompatibility complex class III
 RT region and its comparison to mouse.";
 RL Genome Res. 13:2621-2636(2003).
 DR EMBL; AF109719; AAC82479.1; -; Genomic DNA.
 DR HSSP; Q9SHE7; 1BT0.
 DR Ensembl; ENSMUSG00000024392; Mus musculus.
 DR MGI; MGI:1919439; Bat3.
 DR GO; GO:0006464; P:protein modification; IEA.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS00299; UBIQUITIN_1; 1.
 DR PROSITE; PS00053; UBIQUITIN_2; 1.
 SQ SEQUENCE 1154 AA; 121037 MW; 7F3FD14DF5AC1211 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1154;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPFLMSWLT 13
 DB 928 VNPSLVSWLT 937

RESULT 46
 Q5SQ39 HUMAN
 ID Q5SQ39_HUMAN PRELIMINARY; PRT; 1162 AA.
 AC Q5SQ39;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE HLA-B associated transcript 3.
 GN Name=BAT3; ORFNames=DAQB-195H10.3-018;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Wood J.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL805934; CA118506.1; -; Genomic DNA.
 DR GO; GO:0006464; P:protein modification; IEA.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS00299; UBIQUITIN_1; 1.
 DR PROSITE; PS00053; UBIQUITIN_2; 1.
 SQ SEQUENCE 1162 AA; 122341 MW; FB67886ACE6F6432 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1162;
 Best Local Similarity 80.0%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VGPFLMSWLT 13
 DB 936 VNPSLVSWLT 945

RESULT 47
 O95874 HUMAN
 ID O95874_HUMAN PRELIMINARY; PRT; 1229 AA.
 AC O95874;

```

DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656967; DOI=10.1101/gr.1736803;
RA Xie T., Rowen L., Aguado B., Ahearn M.E., Madan A., Qin S.,
RA Campbell R.D., Hood L.;
RT "Analysis of the gene-dense major histocompatibility complex class III
RT region and its comparison to mouse.";
RL Genome Res. 13:2621-2636(2003).
DR EMBL; AF129756; AAD18085.1; -; Genomic_DNA.
DR HGSP; Q9GHE7; 1870.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
SQ SEQUENCE 1229 AA; 130200 MW; 2EALF2AFB6DCD221 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLSWLT 13
Db 1003 VNPSLSWLT 1012

RESULT 48
Q5SQ35 HUMAN PRELIMINARY; PRT; 1229 AA.
AC Q5SQ35;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3 (Fragment).
GN Name=BAT3; ORFNames=DAQB-195H10.3-014;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CA118510.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 1229 AA; 130200 MW; 2EALF2AFB6DCD221 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLSWLT 13
Db 1003 VNPSLSWLT 1012

RESULT 48
Q5SQ35 HUMAN PRELIMINARY; PRT; 1229 AA.
AC Q5SQ35;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE HLA-B associated transcript 3 (Fragment).
GN Name=BAT3; ORFNames=DAQB-195H10.3-014;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL805934; CA118510.1; -; Genomic_DNA.
DR GO; GO:0006464; P:protein modification; IEA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 1229 AA; 130200 MW; 2EALF2AFB6DCD221 CRC64;

Query Match 54.3%; Score 44; DB 2; Length 1229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VGPSLSWLT 13
Db 1003 VNPSLSWLT 1012

```

```

RESULT 49
Q4HCM7_9DEIO PRELIMINARY; PRT; 108 AA.
AC Q4HCM7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DgeoDRAFT_2030;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.,
RA "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000001; EAL83924.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11647 MW; 76EC53C1650B1858 CRC64;

Query Match 53.1%; Score 43; DB 2; Length 108;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GPSLSMWLTC 14
Db 48 GPAVSGWLTC 57

RESULT 50
Q9Z537_STRCO PRELIMINARY; PRT; 224 AA.
AC Q9Z537;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SCO6135; ORFNames=SC9B2.22c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces

```

```
RT coelcolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL539126; CAA22802.1; -; Genomic_DNA.
DR PIR; T35918; T35918.
KW Complete proteome.
SQ SEQUENCE 224 AA; 23633 MW; ACFAAA6D2F502F8C CRC64;

Query Match      53.1%; Score 43; DB 2; Length 224;
Best Local Similarity 72.7%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 ELVGPSLMSWL 12
      |||||
Db      162 ELVGPELWPWL 172

Search completed: May 12, 2006, 10:50:20
Job time : 134.077 secs
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:38:31 ; Search time 21.8462 Seconds

(without alignments)
79.277 Million cell updates/sec

Title: US-10-632-388-80

Perfect score: 97

Sequence: 1 TTKGPTLRQWLKREHTS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

PIR 80:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	50.5	122	2 A10380	conserved hypotet
2	47	48.5	327	2 B71900	hypothetical prote
3	45	46.4	186	2 A90167	adenylate cyclase,
4	44	45.4	161	2 T06826	beta-fructofuranos
5	44	45.4	1019	2 T11560	pol polyprotein -
6	44	45.4	1058	2 S08436	pol polyprotein -
7	43	44.3	63	2 T30614	hypothetical prote
8	43	44.3	216	2 S35151	photosystem I chai
9	43	44.3	217	2 S46354	pol polyprotein -
10	43	44.3	233	2 A83862	initiation of chro
11	43	44.3	241	2 S07740	hypothetical prote
12	43	44.3	269	2 S73999	hypothetical prote
13	43	44.3	656	2 S30484	pol polyprotein -
14	43	44.3	656	2 S30483	pol polyprotein -
15	43	44.3	1583	2 S59644	sister chromatid c
16	42.5	43.8	275	2 AC0189	probable exported
17	42	43.3	114	2 E71171	hypothetical prote
18	42	43.3	171	2 T20567	hypothetical prote
19	42	43.3	416	2 D72456	probable glutamyl-
20	42	43.3	438	2 S71157	cytochrome c bioge
21	42	43.3	519	2 C96160	hypothetical prote
22	42	43.3	581	2 T12095	beta-fructofuranos
23	42	43.3	617	2 S75447	proline-RNA ligas
24	42	43.3	1061	1 GNLJG4	HIV-1 retropepsin
25	42	43.3	2108	2 H70819	probable polyketid
26	41.5	42.8	877	2 T03098	p97 protein - Toxo
27	41	42.3	70	2 T06920	ribosomal protein
28	41	42.3	153	2 A37524	hypothetical prote
29	41	42.3	178	2 AB2743	hypothetical prote

30	41	42.3	188	2 C82863	recombinase xPa001
31	41	42.3	203	2 T40206	hypothetical prote
32	41	42.3	238	2 S76936	hypothetical prote
33	41	42.3	257	2 E70429	tRNA guanine-N1 me
34	41	42.3	264	2 AG2095	hypothetical prote
35	41	42.3	306	2 D70601	UTP-glucose-1-phos
36	41	42.3	380	2 B47029	methylase Llapi -
37	41	42.3	397	2 G69449	tryptophan synthas
38	41	42.3	419	1 S29127	carboxypeptidase A
39	41	42.3	472	2 T70876	probable papA3 pro
40	41	42.3	622	2 S35122	site-specific DNA-
41	41	42.3	1039	2 S46347	pol polyprotein -
42	41	42.3	1123	2 T51517	telomerase reverse
43	41	42.3	1140	2 S73786	hypothetical prote
44	41	42.3	1191	2 T31091	hypothetical prote
45	41	42.3	3345	2 T13423	hypothetical prote
46	41	41.2	158	2 D72305	hypothetical prote
47	40	41.2	185	2 T49611	hypothetical prote
48	40	41.2	200	2 T23485	hypothetical prote
49	40	41.2	207	2 T37464	probable glutathio
50	40	41.2	236	1 T24395	dihydropteridine r
51	40	41.2	242	2 C70895	hypothetical prote
52	40	41.2	249	2 E87575	ABC transporter, A
53	40	41.2	262	2 B83126	probable transcrip
54	40	41.2	306	2 T45453	UTP-glucose-1-phos
55	40	41.2	336	2 E72389	hypothetical prote
56	40	41.2	344	2 E84377	protein export lim
57	40	41.2	349	2 B97912	phosphoglucose 4,6-de
58	40	41.2	395	2 S40171	phosphoprotein pho
59	40	41.2	419	1 CPRTA	carboxypeptidase A
60	40	41.2	505	2 T19971	hypothetical prote
61	40	41.2	506	2 T19973	hypothetical prote
62	40	41.2	521	2 T01923	hypothetical prote
63	40	41.2	527	2 B64633	hypothetical prote
64	40	41.2	612	2 JQ1346	glucan 1,4-alpha-g
65	40	41.2	978	2 B89971	conserved hypotet
66	40	41.2	1009	2 S28081	pol polyprotein -
67	40	41.2	1009	2 S44621	C50C3.2 protein -
68	40	41.2	1034	1 GNLJCA	HIV-1 retropepsin
69	40	41.2	1035	1 GNLJGG	HIV-1 retropepsin
70	40	41.2	1036	1 GNLJG2	HIV-1 retropepsin
71	40	41.2	1040	2 T08190	hypothetical prote
72	40	41.2	1054	1 GNLJG5	HIV-1 retropepsin
73	40	41.2	1055	1 GNLJST	HIV-1 retropepsin
74	40	41.2	1056	2 S53092	pol polyprotein -
75	40	41.2	1056	1 GNLJG3	HIV-1 retropepsin
76	40	41.2	1712	1 CGH2B	collagen alpha 2(I
77	40	41.2	2609	2 T40399	probable transport
78	39.5	40.7	642	2 A83268	probable soluble l
79	39.5	40.7	722	2 T37970	probable G2-specif
80	39.5	40.7	722	2 T34393	hypothetical prote
81	39	40.2	1758	2 T36141	probable nicotinam
82	39	40.2	220	2 AC0318	probable nicotinam
83	39	40.2	240	2 S73922	uracil DNA glycosy
84	39	40.2	240	2 AE1145	phosphoribosylform
85	39	40.2	314	2 H90638	hypothetical prote
86	39	40.2	318	2 S17197	nitrate reductase
87	39	40.2	320	2 H85489	hypothetical prote
88	39	40.2	327	2 E82277	probable transposa
89	39	40.2	331	2 B48445	glyceraldehyde-3-p
90	39	40.2	336	2 A47542	short-chain alchoh
91	39	40.2	349	2 F86649	dtDP-glucose 4,6-d
92	39	40.2	373	2 D64729	probable activator
93	39	40.2	410	2 H86290	hypothetical prote
94	39	40.2	533	2 G83658	hypothetical prote
95	39	40.2	567	2 S69778	adhesin AP65-1 pre
96	39	40.2	571	2 S48384	DNA43 protein - ye
97	39	40.2	615	2 T15575	hypothetical prote
98	39	40.2	644	1 W1W158	E1 protein - human
99	39	40.2	721	2 A39707	erythrocyte membra
100	39	40.2	721	2 A33319	microtubule-associ
101	39	40.2	859	2 G86242	hypothetical prote
102	39	40.2	877	2 S72541	nitrate reductase

103	39	40.2	989	2	T15576	hypothetical prote	176	37	38.1	180	2	T09063	hypothetical prote
104	39	40.2	1055	1	A37205	leukotoxin A - Act	177	37	38.1	209	2	A83849	alpha-ribazole-5'-
105	39	40.2	1339	2	T40245	probable transcrip	178	37	38.1	214	2	A12371	hypothetical prote
106	39	40.2	1544	2	E59431	phosphoinositide-b	179	37	38.1	232	2	S40165	glutathione transf
107	38.5	39.7	256	2	D90026	hypothetical prote	180	37	38.1	241	2	G90190	conserved hypotet
108	38.5	39.7	416	1	JC4952	transcription init	181	37	38.1	255	2	A80537	hypothetical prote
109	38.5	39.7	481	2	T49321	related to heterok	182	37	38.1	265	2	A56838	uroporphyrinogen-I
110	38.5	39.7	529	2	A10986	probable membrane	183	37	38.1	268	2	D33465	lic-1 protein D -
111	38.5	39.7	1073	2	I51055	recombination acti	184	37	38.1	273	2	A80141	conserved hypotet
112	38.5	39.7	2476	2	T34022	zonadhesin - pig	185	37	38.1	276	2	D82169	conserved hypotet
113	38	39.2	104	2	A96002	hypothetical prote	186	37	38.1	319	2	S59407	ribosomal protein
114	38	39.2	134	2	B75468	hypothetical prote	187	37	38.1	326	2	C24430	glyceralddehyde-3-p
115	38	39.2	182	2	H95917	hypothetical prote	188	37	38.1	327	2	H82736	hypothetical prote
116	38	39.2	194	2	S70990	hypothetical prote	189	37	38.1	330	2	C87036	probable conserved
117	38	39.2	241	2	T23428	hypothetical prote	190	37	38.1	336	1	DEPZG	glyceralddehyde-3-p
118	38	39.2	250	2	T14548	beta-fructofuranos	191	37	38.1	337	2	A35080	glyceralddehyde-3-p
119	38	39.2	254	1	G69592	branched-chain am	192	37	38.1	338	1	DEIS3C	glyceralddehyde-3-p
120	38	39.2	256	1	MFNZ	matrix protein - h	193	37	38.1	338	2	JQ1287	glyceralddehyde-3-p
121	38	39.2	257	2	F72116	conserved hypotet	194	37	38.1	339	2	AG0799	probable transcrip
122	38	39.2	257	2	E86507	hypothetical prote	195	37	38.1	340	2	AD0701	tetrathionate redu
123	38	39.2	260	2	S72748	Bil17F3136 prote	196	37	38.1	352	2	B83000	dtpp-D-glucose 4,6
124	38	39.2	281	2	T04858	hypothetical prote	197	37	38.1	361	2	AF0767	dtpp-glucose 4,6-d
125	38	39.2	311	1	RGCK	regulatory protein	198	37	38.1	361	2	S78542	dtpp-glucose 4,6-de
126	38	39.2	311	2	AH0867	transcription acti	199	37	38.1	361	2	H64969	dtpp-glucose 4,6-de
127	38	39.2	311	2	C85936	positive regulator	200	37	38.1	361	2	S15299	dtpp-glucose 4,6-de
128	38	39.2	311	2	H91090	positive regulator	201	37	38.1	361	2	S41534	dtpp-glucose 4,6-de
129	38	39.2	321	2	T42591	gene 48 protein -	202	37	38.1	368	2	F90152	sun (fmu) protein
130	38	39.2	327	2	A64613	conserved hypotet	203	37	38.1	370	2	AB2578	membrane-bound lyc
131	38	39.2	331	2	A72514	hypothetical prote	204	37	38.1	370	2	A97360	outer membrane lip
132	38	39.2	335	2	E89819	hypothetical prote	205	37	38.1	373	2	AB0604	probable oxidoredu
133	38	39.2	342	1	S64042	porphobilinogen sy	206	37	38.1	379	2	JQ2272	formate dehydrogen
134	38	39.2	346	2	T10173	sterol 24-C-methyl	207	37	38.1	396	2	A2835	sarcosine oxidase
135	38	39.2	356	2	C87699	dtpp-D-glucose-4,6	208	37	38.1	407	2	A86298	hypothetical prote
136	38	39.2	357	2	C89880	hypothetical prote	209	37	38.1	409	2	T47118	thiamine pyridinyl
137	38	39.2	367	2	T06780	probable sterol 24	210	37	38.1	413	2	A97613	sarcosine oxidase
138	38	39.2	368	2	G65119	hypothetical 40.4	211	37	38.1	420	2	C83246	probable binding p
139	38	39.2	368	2	H95203	GTP-binding protei	212	37	38.1	422	2	F96826	hypothetical prote
140	38	39.2	368	2	H98070	conserved hypotet	213	37	38.1	428	2	B71403	hypothetical prote
141	38	39.2	375	2	A47117	formate dehydrogen	214	37	38.1	433	2	S51837	glyceralddehyde-3-p
142	38	39.2	393	2	B85992	probable transport	215	37	38.1	433	2	S51836	glyceralddehyde-3-p
143	38	39.2	399	2	F91146	probable transport	216	37	38.1	437	2	T45481	O-acetylhomoserine
144	38	39.2	438	2	E97342	high affinity gluc	217	37	38.1	469	2	A99656	hypothetical prote
145	38	39.2	473	2	E84853	hypothetical prote	218	37	38.1	490	2	T38088	protoporphyrinogen
146	38	39.2	533	2	S43526	amidophosphoribos	219	37	38.1	491	2	G89860	hypothetical prote
147	38	39.2	610	2	S19461	probable membrane	220	37	38.1	499	2	B85507	hypothetical prote
148	38	39.2	623	2	D71435	hypothetical prote	221	37	38.1	502	2	S43123	protein-tyrosine-p
149	38	39.2	632	2	A11607	D-1-deoxyxylulose	222	37	38.1	525	2	A85362	leucyl aminopeptid
150	38	39.2	636	2	A12552	transposase alr801	223	37	38.1	525	2	C69794	glutamate synthase
151	38	39.2	649	2	S74823	N-acetylmuramoyl-L	224	37	38.1	536	1	SVCEB	2,3-dihydroxybenzo
152	38	39.2	670	2	T02092	beta-fructofuranos	225	37	38.1	536	2	E85558	2,3-dihydroxybenzo
153	38	39.2	735	2	T41187	hypothetical prote	226	37	38.1	536	2	A93708	hypothetical prote
154	38	39.2	807	2	T24110	hypothetical prote	227	37	38.1	540	2	AC0063	hypothetical prote
155	38	39.2	952	2	T28139	TiN6.4 protein - A	228	37	38.1	547	2	T48551	fructosidase-like
156	38	39.2	1004	2	T30641	probable DNA-direc	229	37	38.1	567	2	S69779	adhesin AP65-2 pre
157	38	39.2	1045	2	S23570	pol polyprotein ho	230	37	38.1	568	2	E90364	hypothetical prote
158	38	39.2	1102	2	S55100	hypothetical prote	231	37	38.1	580	2	C82551	phage-related term
159	38	39.2	1114	2	B86423	hypothetical prote	232	37	38.1	586	2	B84271	glutamyl-tRNA sync
160	38	39.2	1123	2	T28139	PK4 protein kinase	233	37	38.1	631	2	B87250	dnak protein (impo
161	38	39.2	1124	1	GNLJFP	HIV-1 retroprotein	234	37	38.1	664	2	G89894	protein kinase [im
162	38	39.2	1124	2	S23820	pol polyprotein -	235	37	38.1	752	2	T35244	UvrA-like ABC tran
163	38	39.2	1199	2	A89813	glutamate synthase	236	37	38.1	775	2	AB3402	topoisomerase IV c
164	38	39.2	1522	2	C96578	hypothetical prote	237	37	38.1	816	2	A71006	hypothetical prote
165	38	39.2	2078	2	T25400	hypothetical prote	238	37	38.1	860	2	C86203	hypothetical prote
166	38	39.2	2285	1	G02434	DNA-directed DNA p	239	37	38.1	907	2	AB1885	hypothetical prote
167	37.5	38.7	142	2	H81223	hypothetical prote	240	37	38.1	984	2	T48216	hypothetical prote
168	37.5	38.7	291	2	H81179	membrane protein N	241	37	38.1	990	1	G46335	env polyprotein pr
169	37.5	38.7	291	2	F81923	probable ABC-trans	242	37	38.1	1032	2	S12153	pol polyprotein -
170	37.5	38.7	296	2	AG0147	probable membrane	243	37	38.1	1204	2	T18812	hypothetical prote
171	37.5	38.7	513	1	S50216	translation initia	244	37	38.1	1254	2	A54818	myosin-VI [similar
172	37.5	38.7	753	2	AH0097	probable biotin su	245	37	38.1	1333	2	A37488	Ras guanine nucleo
173	37	38.1	77	2	AB3468	hypothetical prote	246	37	38.1	1336	2	S25716	glutamate synthase
174	37	38.1	87	2	S42171	cytochrome-c oxida	247	37	38.1	1420	2	A32869	apolipoprotein(a)
175	37	38.1	139	2	S77504	yef21 protein - Sy	248	37	38.1	1514	2	T34869	

249	37	38.1	2183	1	G48556	genome polyprotein	322	36	37.1	452	2	T05699	hypothetical prote
250	37	38.1	2183	1	ZLNZMV	genome polyprotein	323	36	37.1	464	2	G70362	hypothetical prote
251	37	38.1	2319	2	A47004	coagulation factor	324	36	37.1	469	2	AD1926	hypothetical prote
252	37	38.1	2329	2	S44625	C50C3.6 protein -	325	36	37.1	471	2	B38637	Ras inhibitor (clo
253	37	38.1	2363	2	T38841	probable pre-mRNA	326	36	37.1	489	2	A07802	lysine-specific pe
254	37	38.1	2427	2	T16613	hypothetical prote	327	36	37.1	491	2	H83658	hypothetical prote
255	37	38.1	2605	2	T18552	saframycin Mx1 syn	328	36	37.1	505	2	H85361	leucyl aminopeptid
256	37	38.1	3430	1	GNWVWV	genome polyprotein	329	36	37.1	507	2	C82901	conserved hypotet
257	37	38.1	3433	1	GNWVWV	genome polyprotein	330	36	37.1	523	2	T18700	hypothetical prote
258	37	38.1	4485	2	T08044	dyein gamma heavy	331	36	37.1	535	1	S76953	protein kinase (EC
259	36.5	37.6	172	2	T36107	probable serine/ar	332	36	37.1	587	2	S36231	beta-fructofuranos
260	36.5	37.6	341	2	C70932	probable dehydroge	333	36	37.1	590	2	T02096	beta-fructofuranos
261	36.5	37.6	515	1	A59309	interferon-inducib	334	36	37.1	593	2	T01575	beta-fructofuranos
262	36.5	37.6	746	2	F95890	probable aldehyde e	335	36	37.1	648	1	W1WLC1	E1 protein - pygmy
263	36.5	37.6	830	2	T37973	rad16 nucleotide	336	36	37.1	664	2	H83962	serine/threonine p
264	36.5	37.6	870	2	A96637	hypothetical prote	337	36	37.1	666	2	T24170	hypothetical prote
265	36.5	37.6	1876	2	T28627	vitellogenin - Rip	338	36	37.1	701	1	FOFV1R	gag polyprotein -
266	36	37.1	64	2	S41873	transcription fact	339	36	37.1	701	2	S11454	gag polyprotein -
267	36	37.1	86	2	AE2441	hypothetical prote	340	36	37.1	701	2	F48613	gag polyprotein -
268	36	37.1	107	2	S52710	probable neuropept	341	36	37.1	701	2	D48613	gag polyprotein -
269	36	37.1	111	2	F75052	hypothetical prote	342	36	37.1	701	2	S35430	gag polyprotein -
270	36	37.1	114	2	E86353	protein F2E2.12 [i	343	36	37.1	704	2	T01772	hypothetical prote
271	36	37.1	123	1	LAHO	alpha-lactalbumin	344	36	37.1	719	2	B95325	conserved hypotet
272	36	37.1	123	2	S28933	alpha-lactalbumin	345	36	37.1	723	2	G84507	hypothetical prote
273	36	37.1	124	2	F72747	hypothetical prote	346	36	37.1	727	2	AC0076	beta-glucosidase (
274	36	37.1	130	2	I83571	probable membrane	347	36	37.1	760	2	T34414	hypothetical prote
275	36	37.1	130	2	C85701	hypothetical prote	348	36	37.1	765	2	A69440	conserved hypotet
276	36	37.1	130	2	F90843	hypothetical prote	349	36	37.1	786	2	G88065	protein T16A1.2 [i
277	36	37.1	131	2	S74539	hypothetical prote	350	36	37.1	799	2	T02456	protein kinase hom
278	36	37.1	151	1	S45108	hypothetical prote	351	36	37.1	810	1	PLHU	plasmin (EC 3.4.21
279	36	37.1	160	2	T09947	hypothetical prote	352	36	37.1	827	2	S29955	surface glycoprote
280	36	37.1	169	2	S26011	hypothetical prote	353	36	37.1	855	1	JQ2004	env polyprote
281	36	37.1	172	2	AD3606	molybdopter	354	36	37.1	882	2	G83018	pyruvate dehydroge
282	36	37.1	188	2	B25317	gag polyprote	355	36	37.1	896	2	A41273	glutamate receptor
283	36	37.1	203	2	H96525	probable terpene c	356	36	37.1	906	2	A40222	glutamate receptor
284	36	37.1	218	2	JC7220	nuclear protein SR	357	36	37.1	906	2	S38723	glutamate receptor
285	36	37.1	218	2	F70574	hypothetical prote	358	36	37.1	906	2	S25852	glutamate receptor
286	36	37.1	219	2	B83889	hypothetical prote	359	36	37.1	908	2	T22376	hypothetical prote
287	36	37.1	221	2	E64305	conserved hypotet	360	36	37.1	980	2	T27342	hypothetical prote
288	36	37.1	229	2	JC7219	nuclear protein SR	361	36	37.1	1039	2	G83748	alpha-mannosidase
289	36	37.1	231	2	S76204	hypothetical prote	362	36	37.1	1068	1	A43322	1-phosphatidylinos
290	36	37.1	239	2	A25317	gag polyprote	363	36	37.1	1068	1	I38110	1-phosphatidylinos
291	36	37.1	252	2	S06567	finger protein (cl	364	36	37.1	1079	1	TVFVMI	gag-Rml1-env polyp
292	36	37.1	254	2	S35743	pl9 protein - avia	365	36	37.1	1089	2	T31583	hypothetical prote
293	36	37.1	259	2	AB3572	succinoglycan bios	366	36	37.1	1124	2	B45557	HIV-1 retropepsin
294	36	37.1	267	2	I40327	haif protein - Bord	367	36	37.1	1132	1	QSBL	host specificity p
295	36	37.1	279	2	E83456	hypothetical prote	368	36	37.1	1132	2	H90834	host specificity p
296	36	37.1	279	1	FOFVFR	gag polyprote	369	36	37.1	1137	2	B90734	probable host spec
297	36	37.1	309	1	S37028	exom protein - Rhi	370	36	37.1	1138	2	D85584	probable tail comp
298	36	37.1	309	2	S39857	succinoglycan bios	371	36	37.1	1190	2	T38636	tat binding homolo
299	36	37.1	309	2	C95977	glucosyltransferas	372	36	37.1	1238	1	A40185	virulence protein
300	36	37.1	315	2	H97169	carbamoylphosphate	373	36	37.1	1238	1	S17944	virulence sensor p
301	36	37.1	322	2	B81336	pseudouridylylate sy	374	36	37.1	1257	2	S28764	neurocan precursor
302	36	37.1	323	2	G82135	probable lytic mur	375	36	37.1	1271	2	T43269	microcystatin synthe
303	36	37.1	333	2	C81245	APC transporter, p	376	36	37.1	1271	2	A48613	gag/pol polyprotei
304	36	37.1	333	2	B82024	probable thiamin-b	377	36	37.1	1603	2	A48613	S-layer protein -
305	36	37.1	337	2	F91191	probable LPS biosy	378	36	37.1	1616	2	T17884	probable polyketid
306	36	37.1	337	2	G86038	probable LPS biosy	379	36	37.1	1733	2	D70887	polyketide synthas
307	36	37.1	350	2	B38535	A/G-specific adeni	380	36	37.1	1784	2	E86921	transcription init
308	36	37.1	350	2	H85953	adenine glycosylas	381	36	37.1	2068	2	A47371	hypothetical prote
309	36	37.1	350	2	E91108	adenine glycosylas	382	36	37.1	2962	2	T19756	Ran-binding protei
310	36	37.1	372	2	T18697	hypothetical prote	383	36	37.1	3234	1	S58884	elastic titin - hu
311	36	37.1	373	2	F63020	UDP-glucose-heptos	384	36	37.1	7962	2	I38346	Acl protein - mous
312	36	37.1	376	1	F64705	conserved hypotet	385	35.5	36.6	124	2	A54773	probable transposa
313	36	37.1	376	2	F71815	hypothetical prote	386	35.5	36.6	221	2	B60634	anthranilate dioxy
314	36	37.1	382	2	G95896	probable transcrip	387	35.5	36.6	340	2	A83332	two-component sens
315	36	37.1	388	2	F89773	hypothetical prote	388	35.5	36.6	377	2	E84103	protein kinase - B
316	36	37.1	402	2	AD3370	NAD(PAD)-utililin	389	35.5	36.6	386	2	I39814	protein kinase - B
317	36	37.1	409	2	E83179	hypothetical prote	390	35.5	36.6	539	1	D69615	2,3-dihydroxybenzo
318	36	37.1	419	2	E83904	hypothetical prote	391	35.5	36.6	547	2	G83116	pyochelin biosynth
319	36	37.1	429	2	S20050	transcription fact	392	35.5	36.6	558	2	S13579	oligo-1,6-glucosid
320	36	37.1	445	2	H97075	4-aminobutyrate am	393	35.5	36.6	645	2	G72256	hydrogenase (EC 1.
321	36	37.1	451	2	T49341	cytochrome P450 re	394	35.5	36.6	1896	2	T08851	Down syndrome cell

395	35.5	36.6	3429	2	T13853	hypothetical prote	468	35	36.1	433	2	S55968	probable membrane
396	35	36.1	32	2	F46376	1-aminocyclopropan	469	35	36.1	437	2	G81117	conserved hypothet
397	35	36.1	85	2	F75291	hypothetical prote	470	35	36.1	437	2	G81902	probable oxidoredu
398	35	36.1	99	2	S47084	lipid transfer lik	471	35	36.1	451	2	A95379	HYPOHETICAL 50.8
399	35	36.1	119	2	C89582	protein C18A11.1 [472	35	36.1	456	2	T06136	aspartate transami
400	35	36.1	120	2	A97655	hypothetical prote	473	35	36.1	463	2	B84279	hypothetical prote
401	35	36.1	120	2	AG2878	conserved hypothet	474	35	36.1	465	2	I39473	Na+-dependent phos
402	35	36.1	123	2	A26541	alpha-lactalbumin	475	35	36.1	465	2	A56410	sodium/phosphate t
403	35	36.1	159	2	I39605	hetP protein Ana	476	35	36.1	467	2	M48916	sodium phosphate t
404	35	36.1	159	2	AC2158	heterocyst differe	477	35	36.1	471	2	S76290	hypothetical prote
405	35	36.1	160	2	T37000	probable transposa	478	35	36.1	472	2	T20454	hypothetical prote
406	35	36.1	167	2	T16947	hypothetical prote	479	35	36.1	472	2	A13367	multidrug resistan
407	35	36.1	171	2	S30140	DNA-directed RNA p	480	35	36.1	473	2	T06167	beta-fructofuranos
408	35	36.1	173	2	T16685	hypothetical prote	481	35	36.1	499	2	AC2068	cell death suppres
409	35	36.1	189	2	D86909	conserved hypothet	482	35	36.1	500	1	SYHUTD	aspartate-tRNA lig
410	35	36.1	194	2	C70209	conserved hypothet	483	35	36.1	501	1	S1KTDI	aspartate-tRNA lig
411	35	36.1	197	2	T46496	hypothetical prote	484	35	36.1	512	2	B82433	aerobic glycerol-3
412	35	36.1	199	2	T10004	hypothetical prote	485	35	36.1	515	2	B86811	glucan 1,6-alpha-g
413	35	36.1	206	2	T16683	hypothetical prote	486	35	36.1	518	2	B81472	hypothetical prote
414	35	36.1	207	2	C38961	hypothetical prote	487	35	36.1	519	2	S66673	diulfide isomeras
415	35	36.1	207	2	C70856	hypothetical prote	488	35	36.1	529	2	T23431	hypothetical prote
416	35	36.1	212	2	PQ0518	envelope protein -	489	35	36.1	534	2	A29776	glucan 1,4-alpha-g
417	35	36.1	219	2	A46097	GPI-anchor biosynt	490	35	36.1	535	2	T38244	hypothetical prote
418	35	36.1	221	2	T17997	hypothetical prote	491	35	36.1	538	2	H96008	hypothetical prote
419	35	36.1	221	2	F81876	hypothetical prote	492	35	36.1	548	2	H82234	probable glutamate
420	35	36.1	221	2	A82190	hypothetical prote	493	35	36.1	555	2	T06491	beta-fructofuranos
421	35	36.1	224	2	A84697	probable glutathio	494	35	36.1	566	2	T46219	hypothetical prote
422	35	36.1	225	1	Q08E42	BDLF4 protein - hu	495	35	36.1	582	2	T06380	beta-fructofuranos
423	35	36.1	226	2	T29460	hypothetical prote	496	35	36.1	584	2	S37212	probable vetispira
424	35	36.1	239	2	E69189	hypothetical prote	497	35	36.1	598	2	T00509	probable vetispira
425	35	36.1	244	2	T07068	beta-fructofuranos	498	35	36.1	610	2	A82436	penicillin-binding
426	35	36.1	246	2	AH0190	probable oxidoredu	499	35	36.1	615	2	PH0853	methyl-directed mi
427	35	36.1	263	2	UN0258	type II site-speci	500	35	36.1	615	2	B91272	enzyme in methyl-d
428	35	36.1	271	2	C75060	hydrogenase [BC 1	501	35	36.1	615	2	B86113	enzyme in methyl-d
429	35	36.1	272	2	D64155	hypothetical prote	502	35	36.1	617	2	T36294	dihydroxy-acid deh
430	35	36.1	273	2	T51512	hypothetical prote	503	35	36.1	618	2	AG1048	DNA mismatch repai
431	35	36.1	274	2	C83709	hypothetical prote	504	35	36.1	618	2	A33588	mismatch repair pr
432	35	36.1	284	2	A83368	probable thiosulfa	505	35	36.1	625	2	E90236	arginyl-tRNA synth
433	35	36.1	286	2	G70537	hypothetical prote	506	35	36.1	627	2	F81441	probable integral
434	35	36.1	294	2	B95416	probable beta lact	507	35	36.1	632	2	H83106	chemotactic transd
435	35	36.1	296	2	S52354	copper resistance	508	35	36.1	635	2	A87433	hypothetical prote
436	35	36.1	308	2	F86147	hypothetical prote	509	35	36.1	635	2	JC5896	killer cell inhibi
437	35	36.1	326	2	D95197	acetyl xylan ester	510	35	36.1	639	2	J00607	glucan 1,4-alpha-g
438	35	36.1	326	2	A99064	xylan esterase 1 (511	35	36.1	640	1	ALASGR	glucan 1,4-alpha-g
439	35	36.1	331	2	B90173	hypothetical prote	512	35	36.1	640	2	A29166	glucan 1,4-alpha-g
440	35	36.1	333	2	A36925	transcription acti	513	35	36.1	642	2	H81868	probable virulence
441	35	36.1	345	2	A10830	probable RNA methy	514	35	36.1	648	2	T43337	polo-like kinase-1
442	35	36.1	345	2	T08563	dnaJ-related prote	515	35	36.1	651	2	T25953	hypothetical prote
443	35	36.1	346	2	B84602	probable DnaJ prot	516	35	36.1	651	2	A82943	oxidoreductase Atu
444	35	36.1	347	2	AG1817	rod shape-determin	517	35	36.1	667	2	E98339	probable oxidoredu
445	35	36.1	347	2	E64069	yceG protein homol	518	35	36.1	685	1	EFBIS2	suppressor 2 prote
446	35	36.1	355	2	AG1877	hypothetical prote	519	35	36.1	691	2	A54741	erythrocyte membra
447	35	36.1	362	2	A64212	protein serine/thr	520	35	36.1	701	2	S48452	probable membrane
448	35	36.1	365	2	A83220	probable pyruvate	521	35	36.1	705	2	S34271	ribonucleoside-dip
449	35	36.1	368	2	A10462	probable exported	522	35	36.1	709	2	S75212	cong protein - Syn
450	35	36.1	371	2	A88520	41.8k hypothetical	523	35	36.1	712	2	S23650	retrovirus-related
451	35	36.1	388	2	F70933	probable oxidoredu	524	35	36.1	714	2	AG0841	ribonucleoside-dip
452	35	36.1	393	2	AB0035	cystathionine beta	525	35	36.1	725	2	A11544	conserved hypothet
453	35	36.1	399	2	T32126	hypothetical prote	526	35	36.1	757	2	E97230	trehalose/maltose
454	35	36.1	401	2	B75395	aminotransferase,	527	35	36.1	761	2	E84603	probable protein t
455	35	36.1	405	1	T03466	probable exonuclea	528	35	36.1	767	2	T00360	hypothetical prote
456	35	36.1	406	2	A39339	protein C inhibito	529	35	36.1	803	2	E82392	phosphoenolpyruvat
457	35	36.1	407	2	T36140	zinc finger protei	530	35	36.1	805	2	T48840	transducer protein
458	35	36.1	407	2	T36404	probable monooxyge	531	35	36.1	808	1	JQ2043	glycoprotein H pre
459	35	36.1	409	2	T37507	aspartate transami	532	35	36.1	812	1	PLMS	plasmin (EC 3.4.21
460	35	36.1	409	2	S01825	transforming growt	533	35	36.1	837	2	A97557	dimethylglycine de
461	35	36.1	410	2	A41397	transforming growt	534	35	36.1	837	2	AD2777	hypothetical prote
462	35	36.1	410	2	A55706	transforming growt	535	35	36.1	839	2	T20230	replication licens
463	35	36.1	412	2	A34939	transforming growt	536	35	36.1	886	1	JC5085	hypothetical prote
464	35	36.1	412	2	A36169	transforming growt	537	35	36.1	914	2	S18942	hypothetical prote
465	35	36.1	418	2	S46315	aspartate transami	538	35	36.1	921	2	A40597	DNA-directed DNA p
466	35	36.1	423	2	AC1068	probable membrane	539	35	36.1	942	2	C83861	ATP-dependent DNA
467	35	36.1	431	2	G75305	hypothetical prote	540	35	36.1	982	1	VCLJVS	env polyprotein pr

541	35	36.1	983	1	E45390	env polypeptide pr	614	34	35.1	129	2	H95402	hypothetical prote
542	35	36.1	992	2	A83324	probable sensor/re	615	34	35.1	145	2	B91252	hypothetical prote
543	35	36.1	1000	2	T30280	hypothetical prote	616	34	35.1	147	2	S57440	lyd protein - pha
544	35	36.1	1047	2	A12002	hypothetical prote	617	34	35.1	151	2	S63748	HIV-1 retropepsin
545	35	36.1	1080	1	T40587	probable ubiquitin	618	34	35.1	151	2	AB0313	conserved hypotet
546	35	36.1	1086	1	B46335	HIV-1 retropepsin	619	34	35.1	153	2	A96751	hypothetical prote
547	35	36.1	1087	2	JQ1162	Pol protein - Maed	620	34	35.1	154	2	S28181	transcription fact
548	35	36.1	1094	2	F70697	probable arabinosy	621	34	35.1	158	2	A10605	conserved hypotet
549	35	36.1	1101	1	B45390	HIV-1 retropepsin	622	34	35.1	158	2	E64823	hypothetical prote
550	35	36.1	1101	1	GNLVS	HIV-1 retropepsin	623	34	35.1	158	2	E90745	probable sensory t
551	35	36.1	1118	1	SYBCP	carbamoyl-phosphat	624	34	35.1	161	2	A69732	PBSX prophage ORF
552	35	36.1	1148	2	D82091	exodeoxyribonuclea	625	34	35.1	165	2	B87702	ribosomal protein
553	35	36.1	1148	2	T09073	splicing factor S1	626	34	35.1	184	2	H83409	hypothetical prote
554	35	36.1	1191	2	S76414	beta transducin-li	627	34	35.1	188	2	B72450	transcription regu
555	35	36.1	1259	4	GNHUL1	retrovirus-related	628	34	35.1	197	2	G82973	hypothetical prote
556	35	36.1	1275	2	I38588	reverse transcript	629	34	35.1	206	2	T08699	hypothetical prote
557	35	36.1	1275	2	S65824	reverse transcript	630	34	35.1	207	2	G89971	conserved hypotet
558	35	36.1	1275	2	B28096	line-1 protein ORF	631	34	35.1	209	1	XURTMC	conjugative transf
559	35	36.1	1384	2	T26656	hypothetical prote	632	34	35.1	210	2	A69898	hypothetical prote
560	35	36.1	1537	2	F86509	CT147 hypothetical	633	34	35.1	212	2	A39442	hypothetical prote
561	35	36.1	1537	2	C81558	conserved hypotet	634	34	35.1	214	2	A83416	hypothetical prote
562	35	36.1	1537	2	H72112	ct147 hypothetical	635	34	35.1	217	2	T15873	hypothetical prote
563	35	36.1	1657	2	T25421	hypothetical prote	636	34	35.1	219	2	C75637	3-oxoadipate CoA-t
564	35	36.1	1680	2	T01367	hypothetical prote	637	34	35.1	225	2	C83932	hypothetical prote
565	35	36.1	1930	2	F86200	protein F12K11.17	638	34	35.1	228	2	E89923	hypothetical prote
566	35	36.1	2207	2	T42759	Munc13-3 protein -	639	34	35.1	235	2	T19328	hypothetical prote
567	35	36.1	2869	2	T18518	apolipoprotein(a)	640	34	35.1	235	2	G86567	lipote-protein li
568	35	36.1	4377	2	T22864	hypothetical prote	641	34	35.1	235	2	F72056	orodine 5' monop
569	35	36.1	85	2	A55575	hypothetical prote	642	34	35.1	238	2	AH2178	phosphoribosylform
570	34.5	35.6	85	2	B91147	ankyrin 3, long sp	643	34	35.1	240	2	AE1504	beta-fructofuranos
571	34.5	35.6	85	2	F85992	hypothetical prote	644	34	35.1	241	2	T07067	beta-fructofuranos
572	34.5	35.6	85	2	C65120	hypothetical 10.0	645	34	35.1	242	2	C83631	probable carbonic
573	34.5	35.6	119	2	PH1551	Ig H chain V regio	646	34	35.1	245	2	T07071	beta-fructofuranos
574	34.5	35.6	171	2	T48286	hypothetical prote	647	34	35.1	247	1	D24706	modulation protein
575	34.5	35.6	211	2	T03636	GRP-binding protei	648	34	35.1	247	2	PQ0178	glyceraldehyde-3-p
576	34.5	35.6	276	2	D64014	hypothetical prote	649	34	35.1	248	2	PQ0769	phosphoadenyl-su
577	34.5	35.6	291	2	I38167	hypothetical Cpg-1	650	34	35.1	250	2	PQ0768	glycoprotein G - b
578	34.5	35.6	312	2	S34635	superantigen Mtv -	651	34	35.1	251	2	T11549	ATP binding protei
579	34.5	35.6	315	2	S23594	superantigen Mtv(M	652	34	35.1	252	2	T31439	probable cobyric a
580	34.5	35.6	315	2	JH0552	superantigen Mtv3	653	34	35.1	255	1	RHRTT	thyroliberin precu
581	34.5	35.6	315	2	JH0554	superantigen Mtv1	654	34	35.1	255	1	RTMST	thyroliberin precu
582	34.5	35.6	315	2	S33143	superantigen Mtv -	655	34	35.1	256	1	MFNZBR	matrix protein - b
583	34.5	35.6	315	2	JH0551	superantigen Mtv(C	656	34	35.1	256	1	MGNZBR	major surface glyco
584	34.5	35.6	319	1	QQVMTM	superantigen Mtv	657	34	35.1	257	1	MGNZBR	ABC transporter, m
585	34.5	35.6	320	1	QQVMTM	superantigen Mtv	658	34	35.1	261	2	AH2960	nitrate transport
586	34.5	35.6	320	2	S26174	superantigen Mtv17	659	34	35.1	261	2	F98322	hypothetical prote
587	34.5	35.6	320	2	E26387	superantigen Mtv9	660	34	35.1	261	2	F98322	probable transcrip
588	34.5	35.6	320	2	E26795	superantigen Mtv(B	661	34	35.1	264	2	AH2202	bacitracin resist
589	34.5	35.6	322	2	JH0706	superantigen Mtv7	662	34	35.1	265	2	E64128	suppressor protein
590	34.5	35.6	322	2	S24574	superantigen Mtv(S	663	34	35.1	266	2	E71230	hypothetical prote
591	34.5	35.6	324	1	QQVMTM	superantigen - mou	664	34	35.1	271	2	B83682	urease accessory p
592	34.5	35.6	325	2	A49019	superantigen Mtv -	665	34	35.1	272	2	T35231	hypothetical prote
593	34.5	35.6	325	2	S35303	superantigen Mtv -	666	34	35.1	274	2	B83444	probable transcrip
594	34.5	35.6	325	2	F64407	carotenoid biosynt	667	34	35.1	274	2	B75518	hypothetical prote
595	34.5	35.6	359	1	F64407	probable transpos	668	34	35.1	280	2	T43011	hypothetical prote
596	34.5	35.6	369	2	JC4292	hypothetical prote	669	34	35.1	281	2	G72680	ATP synthase F0, s
597	34.5	35.6	472	2	AG2417	probable homeodoma	670	34	35.1	283	2	T23785	hypothetical prote
598	34.5	35.6	482	2	H84539	hypothetical prote	671	34	35.1	286	2	AH2416	hypothetical prote
599	34.5	35.6	524	2	D06021	exoribonuclease II	672	34	35.1	286	2	T35001	probable respirato
600	34.5	35.6	678	2	D82415	homeotic protein H	673	34	35.1	287	2	S43852	neuropeptide Pol-R
601	34.5	35.6	719	2	S25237	melittin - little	674	34	35.1	288	2	S39888	SMR2 protein - Pod
602	34	35.1	26	1	MEHBCF	hypothetical prote	675	34	35.1	288	2	T07730	glyceraldehyde-3-p
603	34	35.1	56	2	AF2385	gene C protein - p	676	34	35.1	295	2	T07730	transcription regu
604	34	35.1	84	1	ZCBP4	hypothetical prote	677	34	35.1	299	2	B83888	hypothetical prote
605	34	35.1	84	2	F84388	hypothetical prote	678	34	35.1	303	2	AI2520	hypothetical prote
606	34	35.1	90	2	AC2180	hypothetical prote	679	34	35.1	306	2	T25954	hypothetical prote
607	34	35.1	92	2	AC2180	glutinin 1 - wheat	680	34	35.1	312	1	QQBE07	US10 protein - hum
608	34	35.1	101	1	EBWT1	cysteine proteinas	681	34	35.1	312	2	F68876	hypothetical prote
609	34	35.1	103	2	T07053	hypothetical prote	682	34	35.1	313	2	AC0098	transcription acti
610	34	35.1	109	2	S69853	hypothetical prote	683	34	35.1	313	2	T49129	hypothetical prote
611	34	35.1	110	2	T47180	hypothetical prote	684	34	35.1	315	2	A40057	transforming growt
612	34	35.1	112	2	S75152	hypothetical prote	685	34	35.1	326	2	S73828	hypothetical prote
613	34	35.1	124	2	T46547	hypothetical prote	686	34	35.1	327	2	AC1615	glycine betaine/ca

687	34	35.1	328	2	AE1252	glycine betaine/ca	760	34	35.1	452	2	D83609	probable glutamine
688	34	35.1	331	2	F82131	glyceraldehyde 3-p	761	34	35.1	454	2	T49300	hypothetical prote
689	34	35.1	333	1	DELOG3	glyceraldehyde-3-p	762	34	35.1	464	2	T17332	hypothetical prote
690	34	35.1	332	1	DELOG3	spore coat protein	763	34	35.1	468	2	T00794	hypothetical prote
691	34	35.1	337	1	DEJUGC	glyceraldehyde-3-p	764	34	35.1	470	2	AD0888	sufl protein limpo
692	34	35.1	337	1	DEJUGC	glyceraldehyde-3-p	765	34	35.1	470	2	E91116	suppressor of ftsi
693	34	35.1	337	1	DESKG	glyceraldehyde-3-p	766	34	35.1	470	2	E85961	suppressor of ftsi
694	34	35.1	337	1	DEUSGM	glyceraldehyde-3-p	767	34	35.1	470	2	G65088	sufl protein precu
695	34	35.1	337	1	DEZMGC	glyceraldehyde-3-p	768	34	35.1	473	2	S04113	nitrogenase (EC 1.
696	34	35.1	337	2	S29814	glyceraldehyde-3-p	769	34	35.1	474	2	C35405	nitrogenase (EC 1.
697	34	35.1	337	2	S26863	glyceraldehyde-3-p	770	34	35.1	476	2	A83235	alkaline phosphata
698	34	35.1	338	1	A47183	hemoglobin precurs	771	34	35.1	477	2	H83588	probable MFS trans
699	34	35.1	338	2	T47218	glyceraldehyde-3-p	772	34	35.1	477	2	T35798	hypothetical prote
700	34	35.1	338	2	JN0452	glyceraldehyde-3-p	773	34	35.1	481	2	T15372	hypothetical prote
701	34	35.1	339	2	G64041	glyceraldehyde-3-p	774	34	35.1	485	2	S52411	ZNF165 protein - h
702	34	35.1	340	2	H83443	glycerol-3-phospha	775	34	35.1	489	1	C84984	lysine-specific pe
703	34	35.1	340	2	T51386	probable protein w	776	34	35.1	489	2	H91009	lysine-specific pe
704	34	35.1	344	2	A95402	probable desaturas	777	34	35.1	489	2	B85854	lysine-specific pe
705	34	35.1	347	2	C90598	hypothetical prote	778	34	35.1	491	2	H84477	probable PttA-like
706	34	35.1	351	1	A48763	transcription fact	779	34	35.1	492	2	D71510	probable oligopep
707	34	35.1	352	2	G84077	hypothetical prote	780	34	35.1	492	2	G81668	peptide ABC transp
708	34	35.1	355	2	T20825	hypothetical prote	781	34	35.1	499	2	C75251	phosphopyruvate hy
709	34	35.1	357	2	T25777	hypothetical prote	782	34	35.1	504	2	A49467	occludin - chicken
710	34	35.1	359	1	LOBP37	DNA ligase (Atp) (783	34	35.1	513	2	AB1921	hypothetical prote
711	34	35.1	359	2	S52650	omega-3 fatty acid	784	34	35.1	515	2	G91271	hypothetical prote
712	34	35.1	360	2	S38570	glyceraldehyde-3-p	785	34	35.1	515	2	G86112	hypothetical prote
713	34	35.1	361	1	F65012	hypothetical prote	786	34	35.1	515	2	S56392	hypothetical prote
714	34	35.1	361	2	A91037	probable peptidase	787	34	35.1	523	2	T30091	hypothetical prote
715	34	35.1	361	2	C85981	probable peptidase	788	34	35.1	524	2	E71881	hypothetical prote
716	34	35.1	362	2	T49528	hypothetical prote	789	34	35.1	525	2	D90048	hypothetical prote
717	34	35.1	365	2	B83829	glycine oxidase (a	790	34	35.1	531	1	SYNCYT	tyrosine-tRNA liga
718	34	35.1	366	2	AH4429	hypothetical prote	791	34	35.1	534	2	T41081	hypothetical prote
719	34	35.1	367	1	I69653	UDPgalactopyranose	792	34	35.1	536	2	AB1983	hypothetical prote
720	34	35.1	368	2	G81289	UDPgalactopyranose	793	34	35.1	544	2	T38469	conserved hypothet
721	34	35.1	377	2	S30088	aciA protein - Eme	794	34	35.1	549	1	S19095	transcription fact
722	34	35.1	379	2	I48133	ubiquinol-cytochro	795	34	35.1	553	2	AH1830	serine/threonine k
723	34	35.1	379	2	I48132	ubiquinol-cytochro	796	34	35.1	557	2	AF3091	hypothetical prote
724	34	35.1	379	2	I48180	ubiquinol-cytochro	797	34	35.1	557	2	D98195	hypothetical prote
725	34	35.1	379	2	I48134	ubiquinol-cytochro	798	34	35.1	559	2	A10135	DNA repair protein
726	34	35.1	382	2	T34789	hypothetical prote	799	34	35.1	562	2	S55516	glucan 1,3-beta-gl
727	34	35.1	384	2	T51423	formate dehydrogen	800	34	35.1	566	2	S41974	RHC21 protein - ye
728	34	35.1	384	2	S34816	nitrogenase cofact	801	34	35.1	566	2	S50979	DNA ligase (Atp) (
729	34	35.1	386	2	P95911	probable hydrocarb	802	34	35.1	568	2	T05218	hypothetical prote
730	34	35.1	389	2	B65096	corrinoid/iron-sul	803	34	35.1	571	2	T33941	hypothetical prote
731	34	35.1	389	2	T33910	probable serine/ch	804	34	35.1	571	2	A82231	transport ATP-bind
732	34	35.1	390	2	A27512	transforming growt	805	34	35.1	580	2	S49308	beta-fructofuranos
733	34	35.1	391	2	S01413	transforming growt	806	34	35.1	580	2	D69645	myo-inositol catab
734	34	35.1	391	2	S01413	transforming growt	807	34	35.1	580	2	H69793	adenine deaminase
735	34	35.1	391	2	E83151	hypothetical prote	808	34	35.1	586	1	JVBPB3	DNA maturase B - p
736	34	35.1	391	2	D83777	hypothetical prote	809	34	35.1	586	1	JVBPB7	DNA maturase B - p
737	34	35.1	399	2	F72417	hypothetical prote	810	34	35.1	592	2	S61503	beta-fructofuranos
738	34	35.1	400	2	S76929	hypothetical prote	811	34	35.1	592	2	S66681	beta-fructofuranos
739	34	35.1	400	2	AF0291	probable drug resi	812	34	35.1	592	2	JQ0991	beta-fructofuranos
740	34	35.1	400	2	T18570	probable malate de	813	34	35.1	603	2	T40528	palmitoyl-protein
741	34	35.1	408	2	T50876	hypothetical membr	814	34	35.1	605	2	T33750	hypothetical prote
742	34	35.1	411	2	AG0045	probable iron-sulf	815	34	35.1	608	2	C96919	glutamine-fructose
743	34	35.1	412	2	D95941	conserved hypothet	816	34	35.1	608	2	T02684	MYB-related transc
744	34	35.1	417	2	AF3448	hypothetical membr	817	34	35.1	618	2	A10171	probable exported
745	34	35.1	417	2	S23675	contrapsin-related	818	34	35.1	620	2	I57937	dopamine transport
746	34	35.1	418	2	S02138	arginine deiminase	819	34	35.1	620	2	A48980	dopamine transport
747	34	35.1	418	2	S76699	hypothetical prote	820	34	35.1	625	1	W1WLR1	E1 protein - rheu
748	34	35.1	427	2	AE1305	conserved hypothet	821	34	35.1	633	2	B70946	NAOH2 dehydrogenas
749	34	35.1	427	2	AE1677	conserved hypothet	822	34	35.1	639	2	T50148	probable ser/thr p
750	34	35.1	430	2	S05481	keratin 18, type I	823	34	35.1	647	1	I58353	LIM protein kinase
751	34	35.1	431	2	AE0551	phosphate regulon	824	34	35.1	647	1	JP0078	LIM protein kinase
752	34	35.1	432	2	AB3060	hypothetical prote	825	34	35.1	662	2	T41442	omnipotent nonsens
753	34	35.1	435	2	T06210	probable glutamate	826	34	35.1	662	2	T51948	omnipotent nonsens
754	34	35.1	438	2	G87337	membrane protein, in	827	34	35.1	662	2	JC7906	sucrose 1F-fructos
755	34	35.1	440	2	A83435	ATP synthase in cy	828	34	35.1	664	1	WMBEBH	infected cell prot
756	34	35.1	443	2	T05132	hypothetical prote	829	34	35.1	668	2	B96999	methyl-accepting c
757	34	35.1	445	2	B75519	probable ubiquinol	830	34	35.1	684	2	F70810	hypothetical prote
758	34	35.1	451	2	B90312	hypothetical prote	831	34	35.1	687	1	B69649	beta-galactosidase
759	34	35.1	451	2	E72398	hypothetical prote	832	34	35.1	687	2	B70515	hypothetical prote

979 33 34.0 105 2 S12183 transcription fact
980 33 34.0 106 2 G90356 hypothetical prote
981 33 34.0 107 2 G85676 unknown protein en
982 33 34.0 107 2 C90817 hypothetical prote
983 33 34.0 120 2 S16445 transcription fact
984 33 34.0 120 2 D90546 proline dipeptidas
985 33 34.0 121 2 T50173 hypothetical prote
986 33 34.0 122 2 AD0327 probable exported
987 33 34.0 125 1 A46315 E4 protein - human
988 33 34.0 129 2 D82708 hypothetical prote
989 33 34.0 130 2 A87666 death on curing pr
990 33 34.0 132 1 S15618 E4 protein - human
991 33 34.0 142 2 AC0928 probable regulator
992 33 34.0 142 2 AF0961 heat shock protein
993 33 34.0 147 2 AF0157 conserved hypothet
994 33 34.0 148 2 AF0067 probable membrane
995 33 34.0 150 2 T34697 probable insertion
996 33 34.0 151 2 T17238 hypothetical prote
997 33 34.0 153 2 C69193 ribosomal protein
998 33 34.0 153 2 AF2392 hypothetical prote
999 33 34.0 154 2 AC0496 heat shock protein
1000 33 34.0 155 2 D75147 lsu ribosomal prot

ALIGNMENTS

RESULT 1
AI0380
conserved hypothetical protein YPO3137 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AI0380
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AI0380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <KUR>
A:Cross-references: UNIPROT:Q8ZC84; UNIPARC:UPI00000DCA48; GB:AL590842; PIDN:CAC92372.1;
C:Genetics:
C:Superfamily: Escherichia coli ybaJ protein

Query Match 50.5%; Score 49; DB 2; Length 122;
Best Local Similarity 53.3%; Pred. No. 1.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSGREH 16
| | | | | | | | | | | | | | | | | | | | | |
Db 92 INDPFLRRWQKTKEH 106

RESULT 2
B71900
hypothetical protein jhp0682 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: B71900
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <ARN>

A:Cross-references: UNIPROT:Q9ZL98; UNIPARC:UPI0000139C5B; GB:AE001500; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
C:Gene: jhp0682
C:Superfamily: conserved hypothetical protein HI0176
Query Match 48.5%; Score 47; DB 2; Length 327;
Best Local Similarity 50.0%; Pred. No. 8;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSGREH 16
: | | | | | | | | | | | | | | | | | | | | | |
Db 103 SVKEPTLVDLWLSQNY 118

RESULT 3
A90167
adenylate cyclase, cyAB-type, probable (cyAB) [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 31-Dec-2004
C:Accession: A90167
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
arrett, R.A.; Ragan, M.A.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A90167
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <KUR>
A:Cross-references: UNIPROT:Q980N7; UNIPARC:UPI00000641DE; GB:AE006641; NID:gl3813390; F

Query Match 46.4%; Score 45; DB 2; Length 186;
Best Local Similarity 55.6%; Pred. No. 9;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSGREHTS 18
| | | | | | | | | | | | | | | | | | | | | |
Db 65 TVKGPKLHSSLKAREIS 82

RESULT 4
T06826
beta-fructofuranosidase (BC 3.2.1.26) II - garden pea (fragment)
N:Alternate names: cell wall invertase II
C:Species: Pisum sativum (garden pea)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06826
R:Buchner, P.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z15838
A:Accession: T06826
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-161 <BUC>
A:Cross-references: UNIPROT:P93490; UNIPARC:UPI00000A9B75; EMBL:Z83339; PIDN:CAB05954.1
C:Superfamily: beta-fructofuranosidase
C:Keywords: glycosidase; hydrolase

Query Match 45.4%; Score 44; DB 2; Length 161;
Best Local Similarity 46.7%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSGREH 16
: | | | | | | | | | | | | | | | | | | | | | |
Db 76 VSDPFLREWIKSPEN 90

RESULT 5

T11560
pol polyprotein - simian immunodeficiency virus SIVsm (strain E543) (fragment)
C:Species: simian immunodeficiency virus SIVsm
A:Variety: strain E543
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11560
R:Hirsch, V.M.; Adger-Johnson, D.; Cambell, B.; Goldstein, S.; Brown, C.; Elkins, W.R.;
J. Virol. 71, 1608-1620, 1997
A:Title: A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficiency virus
A:Reference number: Z17285; MUID:97151152; PMID:899588
A:Accession: T11560
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1019 <HIR>
A:Cross-references: UNIPROT:P89154; UNIPARC:UPI000105470; EMBL:U72748; NID:g1695908; P11560
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: AIDS; immunodeficiency

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 83; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3

QY 3 KGPTLRQWLKRSRE 15
 :|||||:|
Db 184 EGPRLRQWLKRSRE 196

RESULT 6
S08436
pol polyprotein - human immunodeficiency virus type 2 D205 (fragment)
C:Species: human immunodeficiency virus type 2 D205, HIV-2 D205
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S08436
R:Dietrich, U.; Adamski, M.; Kreutz, R.; Seipp, A.; Kuehnle, H.; Ruebsamen-Waigmann, H.
Nature 342, 948-950, 1999
A:Title: A highly divergent HIV-2-related isolate.
A:Accession: S08436
A:Reference number: S08434; MUID:90081881; PMID:2594088
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1058 <DIE>
A:Cross-references: UNIPROT:P15833; UNIPARC:UPI0000131EF6; EMBL:X16109
A:Note: this sequence was submitted to the EMBL Data Library, Aug-1989
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein
C:Keywords: polyprotein

Query Match 45.4%; Score 44; DB 2; Length 1058;
Best Local Similarity 66.7%; Pred. No. 86; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 3

QY 4 GPTLRQWLKRSRE 15
 |||:|||||
Db 222 GPKIRQWLKRSRE 233

RESULT 7
T30614
hypochemical protein 12L - Molluscum contagiosum virus 1
N:Alternate names: MC0012L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T30614
R:Senkevitch, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30614
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

Query Match 44.3%; Score 43; DB 2; Length 63;
Best Local Similarity 41.2%; Pred. No. 5.7; Indels 0; Gaps 0;
Matches 7; Conservative 5; Mismatches 5

QY 2 IKGPTLRQWLKRSREHTS 18
 :|||||:|:
Db 41 VLGETLRTWSRKNTA 57

RESULT 8
S35151
Photosystem I chain XI precursor - spinach
C:Species: Spinacia oleracea (spinach)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S35151; S14446
R:Flieger, K.; Oelmueller, R.; Herrmann, R.G.
Plant Mol. Biol. 22, 703-709, 1993
A:Title: Isolation and characterization of cDNA clones encoding a 18.8 kDa polypeptide,
A:Reference number: S35151; MUID:93344519; PMID:8343606
A:Accession: S35151
A:Molecule type: mRNA
A:Residues: 1-216 <PLI>
A:Cross-references: UNIPROT:Q41385; UNIPARC:UPI000013257B; EMBL:X64445; NID:g396274; P11560
A:Experimental source: seed
R:Ikeuchi, M.; Inoue, Y.
FEBS Lett. 280, 332-334, 1991
A:Title: Two new components of 9 and 14 kDa from spinach photosystem I complex.
A:Reference number: S14316; MUID:91192162; PMID:2013332
A:Accession: S14446
A:Molecule type: protein
A:Residues: 158-175, X, 177-178 <IKE>
A:Cross-references: UNIPARC:UPI000017AF2B
C:Genetics:
A:Gene: psal
C:Function:
A:Description: this protein is a component of photosystem I which catalyzes the light-i
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid; transmembrane protei
F:1-47/Domain: transit peptide (chloroplast) #status predicted <INP>
F:48-216/Product: photosystem I chain XI #status predicted <MAT>
F:131-153/Domain: transmembrane #status predicted <TML>
F:187-209/Domain: transmembrane #status predicted <TM2>

Query Match 44.3%; Score 43; DB 2; Length 216;
Best Local Similarity 52.9%; Pred. No. 22; Indels 7; Gaps 0;
Matches 9; Conservative 1; Mismatches 7

QY 2 IKGPTLRQWLKRSREHTS 18
 :|||||:|:
Db 27 ISGPAURGFPPRRHTS 43

RESULT 9
S46354
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SABD37) (fragment)
C:Species: simian immunodeficiency virus SIVagm
A:Variety: isolate SABD37
C:Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C:Accession: S46354
R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;
EMBO J. 13, 2935-2947, 1994
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African gre
A:Reference number: S46335; MUID:94298785; PMID:8026477
A:Accession: S46354
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-217 <JIN>
A:Cross-references: UNIPARC:UPI000010A53B; EMBL:U04018; NID:g466250; PMID:AAA21512.1; F11560
A:Experimental source: isolate SABD37; sabaeus monkey

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polyprotein
 C:Keywords: polyprotein

Query Match 44.3%; Score 43; DB 2; Length 217;
 Best Local Similarity 66.7%; Pred. No. 22; Mismatches 1; Indels 0; Gaps 0;
 Matches 8; Conservative 1;

QY 4 GPTLRQWLKRSRE 15
 ||||| :
 Db 87 GPLRLQWPLSKS 98
 ||||| :
 ||||| :

RESULT 10
 A83862
 Initiation of chromosome replication dnaD [imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: A83862
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: A83862
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-233 <STO>
 A:Cross-references: UNIPROT:Q9K77; UNIPARC:UPI000003C7C; GB:AP001512; GB:BA000004; NID
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: dnaD

Query Match 44.3%; Score 43; DB 2; Length 233;
 Best Local Similarity 43.8%; Pred. No. 24;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKRSREHT 17
 ||||| :
 Db 144 IEGETLSMWDQDHT 159
 ||||| :
 ||||| :

RESULT 11
 S07740
 Hypothetical protein 8 - Paramaecium tetraurelia mitochondrion
 C:Species: mitochondrion Paramaecium tetraurelia
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
 C:Accession: S07740
 R/Pritchard, A.E.; Seilhamer, J.J.; Mahalingam, R.; Sable, C.L.; Venuti, S.E.; Cummings,
 Nucleic Acids Res. 18, 173-180, 1990
 A:Title: Nucleotide sequence of the mitochondrial genome of Paramaecium.
 A:Reference number: S07725; MUID:90174913; PMID:2308823
 A:Accession: S07740
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-241 <PRI>
 A:Cross-references: UNIPROT:P15609; UNIPARC:UPI000013B890; EMBL:X15917; NID:g13256; PID:
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC6
 A:Start codon: ATT
 C:Superfamily: mitochondrial ribosomal protein S18, paramaecium type
 C:Keywords: mitochondrion

Query Match 44.3%; Score 43; DB 2; Length 241;
 Best Local Similarity 41.2%; Pred. No. 25;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKRSREHTS 18
 ||||| :
 Db 26 VKGPTTEKFLKREYNA 42
 ||||| :
 ||||| :

RESULT 12

S73999
 Hypothetical protein yaaC homolog VxpSPT7_orf269 - Mycoplasma pneumoniae (strain ATCC 29219)
 N:Alternate names: hypothetical protein VxpSPT7_orf269
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C>Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S73999
 R/Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
 A:Reference number: S73327; MUID:97105885; PMID:8948633
 A:Accession: S73999
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-269 <HIM>
 A:Cross-references: UNIPROT:P75587; UNIPARC:UPI00001338E1; EMBL:AF000062; GB:U00089; NID
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: uncharacterized conserved protein HI0963

Query Match 44.3%; Score 43; DB 2; Length 269;
 Best Local Similarity 58.3%; Pred. No. 28;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLK 12
 | : |||||
 Db 124 TLSSSTIRQWLK 135
 | : |||||

RESULT 13

S30484
 pol polyprotein - human immunodeficiency virus type 2
 C:Species: human immunodeficiency virus type 2, HIV-2
 C>Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
 C:Accession: S30484
 R/Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
 submitted to the EMBL Data Library, December 1992
 A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
 A:Reference number: S30460
 A:Accession: S30484
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-656 <GAO>
 A:Cross-references: UNIPARC:UPI00001785D8; EMBL:M87114
 C:Superfamily: pol polyprotein

Query Match 44.3%; Score 43; DB 2; Length 656;
 Best Local Similarity 66.7%; Pred. No. 74;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKRSRE 15
 ||||| :
 Db 31 GPLRLQWPLSKS 42
 ||||| :
 ||||| :

RESULT 14

S30483
 pol polyprotein - human immunodeficiency virus type 2
 C:Species: human immunodeficiency virus type 2, HIV-2
 C>Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 23-Mar-2001
 C:Accession: S30483
 R/Gao, F.; Yue, L.; White, A.T.; Pappas, P.G.; Barchue, J.; Hanson, A.P.; Greene, B.M.;
 submitted to the EMBL Data Library, December 1992
 A:Description: Human infection by genetically diverse SIVSM-related HIV-2 in west Africa
 A:Reference number: S30460
 A:Accession: S30483
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-656 <GAO>
 A:Cross-references: UNIPARC:UPI00001785D7; EMBL:M87111

C:Superfamily: pol polyprotein

Query Match 44.3%; Score 43; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSRE 15
||| ||| |||
Db 31 GPKLRQWPLSKE 42

RESULT 15

S59644

sister chromatid cohesion molecule Mis4p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 14-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C:Accession: T38603; T43392; S59644

R:Devlin, K.; Churcher, C.M.; Barrrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, July 1995

A:Reference number: Z21731

A:Accession: T38603

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1583 <DE2>

A:Cross-references: UNIPROT:Q09725; UNIPARC:UPI000012F161; EMBL:Z50113; NID:G914878; PID

A:Experimental source: strain 972h-; cosmid c31A2

R:Furuya, K.; Takahashi, K.; Yanagida, M.

submitted to the EMBL Data Library, August 1998

A:Description: Faithful anaphase is ensured by Mis4, a sister chromatid cohesion molecule

A:Reference number: Z22478

A:Accession: T43392

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1583 <FUR>

A:Cross-references: UNIPARC:UPI000012F161; EMBL:AB016866; PIDN:CAB19489.1

C:Genetics:

A:Gene: mis4; SPAC31A2.05C

A:Map position: 1

A:Introns: 33/1; 98/2; 543/3; 699/3; 1294/2; 1339/3; 1558/3

Query Match 44.3%; Score 43; DB 2; Length 1583;

Best Local Similarity 46.2%; Pred. No. 1.9e+02;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSREH 16

|||: ||: |:

Db 1483 GPTITGMKKLDH 1495

RESULT 16

AC0189

probable exported protein YPO1551 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AC0189

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0189

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-275 <KUR>

A:Cross-references: UNIPROT:Q8ZFX2; UNIPARC:UPI00000DCBC5; GB:AL590842; PIDN:CAC90374.1;

C:Genetics:

A:Gene: YPO1551

Query Match 43.8%; Score 42.5; DB 2; Length 275;

Best Local Similarity 47.6%; Pred. No. 34;

Matches 10; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

Qy 1 TIKGPT---LROWLKSREHTS 18
|: || | |||
Db 133 TVAGTKLMALEQMLHOLPHTS 153

RESULT 17

E71171

hypothetical protein PH0569 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C:Accession: E71171

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki,

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: E71171

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-114 <KAW>

A:Cross-references: UNIPROT:O58304; UNIPARC:UPI0000062E54; GB:AP000002; NID:G3236129; P

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0569

Query Match 43.3%; Score 42; DB 2; Length 114;

Best Local Similarity 63.6%; Pred. No. 16;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLK 12

: || | |||

Db 47 VKGDTLKVMK 57

RESULT 18

T20567

hypothetical protein F08A10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T20567

R:Kershaw, J.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19293

A:Accession: T20567

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-171 <WIL>

A:Cross-references: UNIPROT:Q9186; UNIPARC:UPI000017B8BA; EMBL:Z75534; PIDN:CAA99825.1

A:Experimental source: clone F08A10

C:Genetics:

A:Gene: CESP.F08A10.1

A:Map position: 1

A:Introns: 26/3; 64/1; 91/3; 118/3

Query Match 43.3%; Score 42; DB 2; Length 171;

Best Local Similarity 57.1%; Pred. No. 24;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 GPTLRQWLKSREHT 17

|||: ||: |:

Db 103 GPSLRPPLNSGNHT 116

RESULT 19

D72456

probable glutamyl-tRNA reductase APE2296 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: D72456

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: D72456

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-416 <KAW>

A:Cross-references: UNIPROT:Q9Y9J2; UNIPARC:UPI000005E2B8; DBJ:AP000064; NID:G5105945;

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2296

C:Superfamily: glutamyl-tRNA reductase

Query Match 43.3%; Score 42; DB 2; Length 416;

Best Local Similarity 57.1%; Pred. No. 65;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRWLKSGRE 15

Db 105 VLQGVPRWLKSGRE 118

RESULT 20

S71157

Cytochrome c biogenesis protein 454 - evening primrose mitochondrion

C:Species: mitochondrion Oenothera lamarckiana (evening primrose)

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004

A:Accession: S71157; S55283; S42984

R:Gruska, I.; Jakabsons, W.; Schuster, W.

Mol. Gen. Genet. 247, 529-536, 1995

A:Title: Oenothera mitochondrialorf454, a gene involved in cytochrome c biogenesis com

A:Reference number: S55283; MUID:95327048; PMID:7603431

A:Accession: S71157

A:Molecule type: mRNA

A:Residues: 1-438 <GRU>

A:Cross-references: UNIPROT:Q35213; UNIPARC:UPI000017CAV5; EMBL:X78036

A>Note: differences are due to RNA editing; premature stop codon is due to RNA editing

A:Accession: S55283

A:Molecule type: DNA

A:Residues: 1-16, 'PR', 19-34, 'P', 36-39, 'SS', 42-48, 'P', 50, 'PS', 53, 'P', 55-108, 'L', 110-130, 'P'

A:Cross-references: UNIPARC:UPI00000964AA; EMBL:X78036; NID:G459536; PIDN:CAA54966.1; PI

C:Genetics:

A:Genome: mitochondrion

A:Introns: 256/2

C:Keywords: mitochondrion; RNA editing

Query Match 43.3%; Score 42; DB 2; Length 438;

Best Local Similarity 50.0%; Pred. No. 68;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 LRQWLKSGREHTS 18

Db 342 LHRVKNRHHN 353

RESULT 21

C86160

hypothetical protein F2D16.3 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

A:Accession: C86160

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86160

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-519 <STO>

A:Cross-references: UNIPROT:Q9SR9Y; UNIPARC:UPI00000A8372; GB:AE005172; NID:G6056405; P:

C:Genetics:

A:Map position: 1

Query Match 43.3%; Score 42; DB 2; Length 519;

Best Local Similarity 40.0%; Pred. No. 82;

Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 2 IKGPTLRWLKSGREH 16

Db 498 IRGHTWVWLKAKQN 512

RESULT 22

T12095

beta-fructofuranosidase (EC 3.2.1.26), cell wall - fava bean

N:Alternate names: cell wall invertase II

C:Species: Vicia faba (fava bean)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

A:Accession: T12095

R:Weber, H.; Borisjuk, L.; Helm, U.; Buchner, P.; Wobus, U.

Plant Cell 7, 1835-1846, 1995

A:Title: Seed coat-associated invertases of Fava bean control both unloading and storage

A:Reference number: Z17416; MUID:96093423; PMID:8535137

A:Accession: T12095

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-581 <WEB>

A:Cross-references: UNIPROT:Q43856; UNIPARC:UPI00000A5748; EMBL:Z35163; NID:G861156; PI

A:Experimental source: strain minor; cultivar Fribo; seed coat; Clone VfcWINV2

C:Genetics:

A:Gene: CWINV2

C:Function:

A:Description: hydrolyzes terminal non-reducing beta-D-fructofuranoside residues in bet

C:Superfamily: beta-fructofuranosidase

C:Keywords: cell wall; glycoprotein; glycosidase; hydrolase

Query Match 43.3%; Score 42; DB 2; Length 581;

Best Local Similarity 58.3%; Pred. No. 93;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PTLROWLKSREH 16

Db 160 PFLREWLKSPEN 171

RESULT 23

S75447

proline-tRNA ligase - Synecocystis sp. (strain PCC 6803)

N:Alternate names: protein sll1425

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75447

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-617 <KAN>

A:Cross-references: UNIPROT:P73942; UNIPARC:UPI0000136614; EMBL:D90911; GB:AB001339; NI

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: pros

A:Start codon: GTG

C:Superfamily: proline-tRNA ligase

```
Query Match 43.3%; Score 42; DB 2; Length 617;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 PTLRWLKSREHTS 18
DB 604 PTLTAWIKAEKTKS 617

RESULT 24
GNLJGA
HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (African green monkey i
N; Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly
C; Species: simian immunodeficiency virus, SIV
C; Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
C; Accession: B30045
R; Fukaawa, M.; Miura, T.; Hasegawa, A.; Morikawa, S.; Teujimoto, H.; Miki, K.; Kitamura
Nature 333, 457-461, 1988
A; Title: Sequence of simian immunodeficiency virus from African green monkey, a new memb
A; Reference number: A30045; MUID:98232906; PMID:3374586
A; Accession: B30045
A; Molecule type: DNA
A; Residues: 1-1061 <FUK>
A; Cross-references: UNIPARC:UPI0000131F18; EMBL:X07805; NID:g1335593
C; Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
C; Genetics:
A; Gene: pol
C; Superfamily: pol polyprotein
C; Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polyprotein; reverse
F;111-210/Product: retropepsin #status predicted <RTP>
F;134/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 43.3%; Score 42; DB 1; Length 1061;
Best Local Similarity 53.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRLWLKSRRE 15
DB 229 RGPCVRQWPLSKRE 241

RESULT 25
H70819
probable polyketide synthase - Mycobacterium tuberculosis (strain H37RV)
C; Species: Mycobacterium tuberculosis
C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C; Accession: H70819
R; Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A; Authors: Squares, R.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A; Reference number: A70500; MUID:98295987; PMID:9634230
A; Accession: H70819
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-2108 <COL>
A; Cross-references: UNIPROT:O53901; UNIPARC:UPI00000D3B08; GB:AL022000; GB:AL123456; NID
A; Experimental source: strain H37RV
C; Genetics:
A; Gene: pks8
C; Superfamily: mycocerosic acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I ho
nase homology; [acyl-carrier-protein] S-malonyltransferase homology
C; Keywords: carrier protein
F;26-434/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F;546-826/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F;1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>
F;1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>
F;2029-2094/Domain: acyl carrier protein homology <ACP1>

Query Match 43.3%; Score 42; DB 2; Length 2108;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Query Match 42.8%; Score 41.5; DB 2; Length 877;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 4 GP---TLRWLKSREH 16
DB 72 GPVNTFQQWLQNEH 87

RESULT 27
T06920
ribosomal protein L28 - Cyanophora paradoxa cyanelle
C; Species: cyanelle Cyanophora paradoxa
C; Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C; Accession: T06920
R; Stirewalt, V. L.; Michalowski, C. B.; Luffelhardt, W.; Bohnert, H. J.; Bryant, D. A.
submitted to the EMBL Data Library, July 1995
A; Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A; Reference number: Z15840
A; Accession: T06920
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-70 <STI>
A; Cross-references: UNIPROT:P48129; UNIPARC:UPI0000133A7D; EMBL:U30821; NID:g1016083; P
A; Experimental source: strain Pringsheim LB555
C; Genetics:
A; Gene: rpl28
A; Genome: cyanelle
C; Superfamily: Escherichia coli ribosomal protein L28
C; Keywords: cyanelle; ribosome

Query Match 42.3%; Score 41; DB 2; Length 70;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRLWLKSRREHT 17
DB 38 IWSPTLRKRLKQVST 53

RESULT 28
A97524
hypothetical protein AGR_C_2500 [imported] - Agrobacterium tumefaciens (strain C58, Ce;
C; Species: Agrobacterium tumefaciens
C; Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C; Accession: A97524
```


Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: G69449
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-397 <KLE>
 A:Cross-references: UNIPROT:028672; UNIPARC:UPI0000056C7D; GB:AE0000782; NID
 C:Genetic: trpB-2
 A:Gene: trpB-2
 C:Function:
 A:Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan
 A:Pathway: tryptophan biosynthesis
 A>Note: cofactor pyridoxal phosphate; last step in pathway
 C:Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate; trypt
 F:12-393/Domain: tryptophan synthase beta chain homology <TRPB>
 F:93/Active site: His #status predicted
 F:94/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
 Query Match 42.3%; Score 41; DB 2; Length 397;
 Best Local Similarity 63.6%; Pred. No. 88;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 7 LRQWLKSRREHT 17
 |||:::|||||
 Db 181 LRQWVESPEHT 191

RESULT 38
 S29127
 carboxypeptidase A (EC 3.4.17.1) CPA1 precursor - human
 N:Alternate names: pancreatic carboxypeptidase A1
 C:Species: Homo sapiens (man)
 C:Date: 25-Feb-1994 #sequence revision 19-Jan-1996 #text change 09-Jul-2004
 C:Accession: S29127; A34205; S08253; S02810; S71394; S02811
 R:Caetanus, L.; Villegas, V.; Pascual, R.; Aviles, F.X.; Wicker-Planquart, C.; Puigserver
 Biochem. J. 287, 299-303, 1992
 A:Title: cDNA cloning and sequence analysis of human pancreatic procarboxypeptidase A1.
 A:Reference number: S29127; MUID:93038569; PMID:1417781
 A:Accession: S29127
 A:Molecule type: mRNA
 A:Residues: 1-419 <CAT>
 A:Cross-references: UNIPROT:PI5085; UNIPARC:UPI000004D332; EMBL:X67318; NID:g35329; PIDN
 R:Stewart, E.A.; Craik, C.S.; Hake, L.; Bowcock, A.M.
 Am. J. Hum. Genet. 46, 795-800, 1990
 A:Title: Human carboxypeptidase A identifies a BgIII RFLP and maps to 7q31-qter.
 A:Reference number: A34205; MUID:90196012; PMID:1969228
 A:Accession: A34205
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 330-396 <STE>
 A:Cross-references: UNIPARC:UPI0000172A5F
 A>Note: the authors translated the codon CTG for residue 391 as Val
 R:Moulard, M.; Michon, T.; Kerfelec, B.; Chapus, C.
 FEBS Lett. 261, 179-183, 1990
 A:Title: Further studies on the human pancreatic binary complexes involving procarboxype
 A:Reference number: S08253; MUID:90169111; PMID:2307232
 A:Accession: S08253
 A:Molecule type: protein
 A:Residues: 17-43; 'XXX', 114-135 <MOU>
 A:Cross-references: UNIPARC:UPI0000172A60; UNIPARC:UPI0000172A61
 R:Pascual, R.; Burgos, F.J.; Salva, M.; Soriano, F.; Mendez, E.; Aviles, F.X.
 Eur. J. Biochem. 179, 609-616, 1989
 A:Title: Purification and properties of five different forms of human procarboxypeptidase
 A:Reference number: S02809; MUID:89153096; PMID:2920728
 A:Accession: S02810
 A:Molecule type: protein
 A:Residues: 17-42 <PAS>
 A:Cross-references: UNIPARC:UPI0000172A62

R:Laethem, R.M.; Blumenkopf, T.A.; Cory, M.; Elwell, L.; Moxham, C.P.; Ray, P.H.; Walton
 Arch. Biochem. Biophys. 332, 8-18, 1996
 A:Title: Expression and characterization of human pancreatic preprocarboxypeptidase A1 e
 A:Reference number: S71394; MUID:96400327; PMID:8806703
 A:Accession: S71394
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-419 <LAE>
 A:Cross-references: UNIPARC:UPI000004D332
 C:Genetic:
 A:Gene: GDB:CPA1; CPA
 A:Cross-references: GDB:120597; OMIM:114850
 A:Map position: 7q32-qter
 C:Superfamily: carboxypeptidase
 C:Keywords: hydrolase; metallo-carboxypeptidase; metalloprotein; protein digestion; zinc
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-110/Domain: activation peptide #status predicted <ACP>
 F:111-419/Product: carboxypeptidase A isozyme 1 #status predicted <MAT>
 F:179-182-306/Binding site: zinc (His, Glu, His) #status predicted
 F:248-277/Diulfide bonds: #status predicted
 F:358-380/Disulfide site: Tyr, Glu #status predicted
 Query Match 42.3%; Score 41; DB 1; Length 419;
 Best Local Similarity 53.8%; Pred. No. 94;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 6 TLQWLKSRREHTS 18
 |||:::|||||
 Db 232 TNRWRKTRSHTA 244

RESULT 39
 F70876
 probable papA3 protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text change 09-Jul-2004
 C:Accession: F70876
 R:Cols, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgates, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: F70876
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-472 <COL>
 A:Cross-references: UNIPROT:O50438; UNIPARC:UPI00000D1152; GB:AL010186; GB:AL123456; NII
 A:Experimental source: strain H37RV
 C:Genetic:
 A:Gene: papA3
 Query Match 42.3%; Score 41; DB 2; Length 472;
 Best Local Similarity 38.9%; Pred. No. 11e+02;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 TIKGPTLRQWLKSRREHTS 18
 |||:::|||||
 Db 224 TVESQVRAWTKFAEGTN 241

RESULT 40
 S35122
 site-specific DNA-methyltransferase (adenine-specific) (EC 2.1.1.72) LlaI - Lactococcus
 N:Alternate names: type II modification methylase LlaI
 C:Species: Lactococcus lactis
 C:Date: 16-Apr-1997 #sequence revision 09-May-1997 #text change 31-Dec-2004
 C:Accession: S35122; S77702; A47029
 R:Hall, C.; Miller, L.A.; Kleenhammer, T.R.
 J. Bacteriol. 173, 4363-4370, 1991
 A:Title: In vivo genetic exchange of a functional domain from a type II A methylase bet
 A:Reference number: A47029; MUID:91294179; PMID:1906061

A:Accession: S35122
A:Molecule type: DNA
A:Residues: 1-622 <HIL>
A:Cross-references: UNIPROT:P35516; UNIPARC:UPI000017AC4E; EMBL:M77136
A:Experimental source: bacteriophage resistance plasmid pIR2030
A:Note: the sequence of residues 469 and 470 is interchanged in the authors' translation
A:Note: sequence extracted from NCBI backbone (NCBIN:41635, NCBIP:41636)
R:Klaenhammer, T. R.
submitted to the EMBL Data Library, November 1994
A:Reference number: S77702
A:Accession: S77702
A:Molecule type: DNA
A:Residues: 1-248, 'G', 250-622 <KLA>
A:Cross-references: UNIPARC:UPI000012F7DD; EMBL:U17233; NID:G639886; PIDN:AAA65073.1; PIR:AA65073.1
C:Genetics:
A:Gene: LlaI
A:Genome: plasmid
A:Superfamily: modification methylase (adenine-specific), M.SteI type
C:Keywords: methyltransferase; restriction modification system; S-adenosylmethionine

Query Match 42.3%; Score 41; DB 2; Length 622;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSGREHT 17
||| ||||| |||||
Db 81 GKTPFQWLNEREXT 94
||| ||||| |||||

RESULT 41
S46347
pol polyprotein - simian immunodeficiency virus SIVagm (isolate SAB-1)
C:Species: simian immunodeficiency virus SIVagm
A:Variety: isolate SAB-1
C:Date: 25-Dec-1994 #sequence_revision 14-Feb-1997 #text_change 26-Aug-1999
C:Accession: S46347
R:Jin, M.J.; Hui, H.; Robertson, D.L.; Mueller, M.C.; Barre-Sinoussi, F.; Hirsch, V.M.;
EMBO J. 13, 2935-2947, 1994
A:Title: Mosaic genome structure of simian immunodeficiency virus from West African green
A:Reference number: S46335; MUID:94298785; PMID:8026477
A:Accession: S46347
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1039 <JIN>
A:Cross-references: UNIPARC:UPI00001096DD; EMBL:U04005; NID:G466229; PIDN:AA21505.1; PIR:AA21505.1
A:Experimental source: isolate SAB-1; sabaesus monkey
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993
A:Note: this reading frame extends between two stop codons and does not begin with a start
C:Genetics:
A:Gene: pol
C:Superfamily: pol polyprotein

Query Match 42.3%; Score 41; DB 2; Length 1039;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSGRE 15
||| ||||| |||||
Db 205 GPRIRQWPLSKS 216
||| ||||| |||||

RESULT 42
T51517
telomerase reverse transcriptase - Arabidopsis thaliana
N:Alternate names: protein F5E19 190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51517
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; New
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51517
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-1123 <SAT>
A:Cross-references: UNIPROT:Q9SPU7; UNIPARC:UPI00000A945F; EMBL:AL391147
A:Experimental source: cultivar Columbia; BAC clone F5E19
C:Genetics:
A:Map position: 5
A:Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2;
A:Note: F5E19_190

Query Match 42.3%; Score 41; DB 2; Length 1123;
Best Local Similarity 58.3%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKS 13
:: || |||||
Db 200 VQPTKRWLSS 211
||| |||||

RESULT 43
S73786
hypothetical protein A19_orf1140 - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S73786
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73786
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1140 <HIM>
A:Cross-references: UNIPROT:P75405; UNIPARC:UPI000013ABAC; EMBL:AE000045; GB:U00089; N1
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3

Query Match 42.3%; Score 41; DB 2; Length 1140;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 GPTLRQWLKSGREH 16
||| ||||| |||||
Db 1116 GVLLHRWRSKRKH 1128
||| |||||

RESULT 44
T31091
hypothetical protein wbbK [imported] - Serratia marcescens
C:Species: Serratia marcescens
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31091
R:Saigi, F.; Climent, N.; Pique, N.; Sanchez, C.; Merino, S.; Rubires, X.; Aguilar, A.;
J. Bacteriol. 181, 1883-1891, 1999
A:Title: Genetic analysis of the Serratia marcescens N28b O4 antigen gene cluster.
A:Reference number: Z20974; MUID:99173913; PMID:10074083
A:Accession: T31091
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1191 <SAI>
A:Cross-references: UNIPROT:O52484; UNIPARC:UPI00000BA53A; EMBL:AF038816; NID:G2828669;
C:Genetics:
A:Gene: wbbK

Query Match 42.3%; Score 41; DB 2; Length 1191;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 PTLRQWLKSGREHTS 18
||| ||||| |||||
Db 491 PELTQWLREAEYA 504
||| |||||

```
RESULT 45
T13423
hypothetical protein 3088.4 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13423
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Accession: T13423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3345 <MUR>
A:Cross-references: UNIPROT:O46074; UNIPARC:UPI00000796D3; EMBL:AL009195; NID:e1355203;
C:Genetics:
A:Cross-references: FlyBase:FBgn0000377
A:Map position: X
A:Introns: 51/3; 159/1; 476/1; 526/1; 1465/1; 1826/3; 1947/3; 2081/1; 2196/3; 3007/3
A:Note: EG:3088.4

Query Match 42.3%; Score 41; DB 2; Length 3345;
Best Local Similarity 58.3%; Pred. No. 9.1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKS 13
Db 2203 VKNPKLEQWLAS 2214
:|:|:|:|:|

RESULT 46
D72305
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72305
R:Nelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72305
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: UNIPROT:Q9X0A5; UNIPARC:UPI00000C1322; GB:AE001763; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1012

Query Match 41.2%; Score 40; DB 2; Length 158;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSRE 15
Db 145 KAETLRKWLNERK 157
|:|:|:|:|:|

RESULT 47
T49611
hypothetical protein B3E4.240 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49611
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49611
A>Status: preliminary
A:Molecule type: DNA
```

```
A:Residues: 1-185 <SCH>
A:Cross-references: UNIPROT:Q9PSV8; UNIPARC:UPI000017B4A4; EMBL:AL355931; GSPDB:GN00116;
A:Experimental source: BAC clone B3E4; strain OR74A
C:Genetics:
A:Gene: NCSP:B3E4.240
A:Map position: 6

Query Match 41.2%; Score 40; DB 2; Length 185;
Best Local Similarity 33.3%; Pred. No. 55;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRE 15
Db 84 TLQGPSAQDWARHRD 98
|:|:|:|:|:|

RESULT 48
T23485
hypothetical protein K08F4.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T23485
R:Hembry, C.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19746
A:Accession: T23485
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-200 <WIL>
A:Cross-references: UNIPARC:UPI00001754D0; EMBL:Z68879; PIDN:CAA93088.1; GSPDB:GN000022;
A:Experimental source: clone K08F4
C:Genetics:
A:Gene: CRSP:K08F4.11
A:Map position: 4
A:Introns: 45/1; 76/1; 111/3
C:Superfamily: glutathione transferase

Query Match 41.2%; Score 40; DB 2; Length 200;
Best Local Similarity 43.8%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
Db 183 IETPKLKSWLAKRPET 198
|:|:|:|:|:|

RESULT 49
T37464
probable glutathione transferase (EC 2.5.1.18) GST3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37464
R:Tawe, W.N.; Eschbach, M.L.; Walter, R.D.; Henkle-Duehrsen, K.
submitted to the EMBL Data Library, June 1997
A:Description: Parquat mediates differential gene expression in C. elegans.
A:Reference number: Z21702
A:Accession: T37464
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-207 <TAW>
A:Cross-references: UNIPROT:O16116; UNIPARC:UPI0000083A0A; EMBL:AF010241; PIDN:AAB65419
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: GST3
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 41.2%; Score 40; DB 2; Length 207;
Best Local Similarity 43.8%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSREHT 17
|:|:|:|:|:|
```

Db 190 IETPKLEWLAKRPET 205

RESULT 50

T24395
dihydropteridine reductase (EC 1.6.99.7) - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T24395
R/Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19886
A/Accession: T24395
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-236 <WIL>
A/Cross-references: UNIPROT:Q9XVJ3; UNIPARC:UPI000007A04C; EMBL:Z81113; PIDD:CA803278.1
A/Experimental source: clone T03F6
C/Genetics:
A/Gene: CESP:T03F6.1
A/Map position: 3
A/Introns: 28/3; 63/2; 104/3; 194/2
C/Function:
A/Description: catalyzes the reduction of quinoxid 7,8-dihydrobiopterin to tetrahydrobiopterin (BH-4) is an essential cofactor for phenylalanine, tyrosine,
A/Note: tetrahydrobiopterin (BH-4) is an essential cofactor for phenylalanine, tyrosine,
C/Superfamily: dihydropteridine reductase
C/Keywords: biopterin; NAD; oxidoreductase
F/2-236/Product: dihydropteridine reductase #status predicted <MAT>
F/5-34/Region: beta-alpha-beta NAD nucleotide-binding fold
F/83,143/Binding site: substrate (Trp, Tyr) #status predicted

Query Match	41.2%	Score 40;	DB 1;	Length 236;
Best Local Similarity	33.3%	Pred. No. 72;		
Matches	6;	Conservative	6;	Mismatches
			6;	Indels
				0;
				Gaps
				0;

Qy 1 TIKGPTLRQWLKSRHTS 18
|:|:|:|:|:
Db 177 TLDTEMNRKWPMPNADHSS 194

Search completed: May 12, 2006, 10:51:37
Job time : 36.9462 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:20:08 ; Search time 127.385 Seconds
(without alignments)
99.694 Million cell updates/sec

Title: US-10-632-388-80

Perfect score: 97

Sequence: 1 TIKGPTLRQWLKSRHTS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	50.5	122	2	Q8ZC84_YERPE
2	49	50.5	122	2	Q6D0R4_YERPS
3	48.5	50.0	718	2	Q5A0W2_EMENI
4	48	49.5	148	2	Q4UK00_RICFE
5	48	49.5	326	2	P95613_RHIGA
6	48	49.5	992	2	Q4P0J6_USTMA
7	47	48.5	327	1	Y745_HELPJ
8	47	48.5	727	2	Q750V6_ASHGO
9	47	48.5	1794	2	Q55Z22_CRYNE
10	47	48.5	1811	2	Q5KNL7_CRYNE
11	46.5	47.9	332	1	GPDA2_MYCPA
12	46	47.4	126	2	Q6YUC3_ORYSA
13	46	47.4	365	2	Q6C114_YARLI
14	46	47.4	402	2	Q82R87_STRAW
15	46	47.4	815	2	Q4KS46_VIRU
16	46	47.4	941	2	Q8QU06_VIRU
17	45	46.4	186	2	Q980N7_SULSO
18	45	46.4	243	2	Q8XR40_RALSO
19	45	46.4	286	2	Q8T462_DROME
20	45	46.4	302	2	Q742B3_MYCPA
21	45	46.4	351	2	Q7V126_HELHP
22	45	46.4	789	2	Q6N3E8_RHOPA
23	45	46.4	1126	2	Q4QH74_LEIMA
24	45	46.4	1361	2	Q5TV94_ANOGA
25	45	46.4	1431	2	Q7PS78_ANOGA
26	44.5	45.9	642	2	Q4KFC8_PSEFP
27	44	45.4	131	2	Q4JW69_CORJK
28	44	45.4	161	2	P93490_PEA
29	44	45.4	298	2	Q5YPT6_NOCFA
30	44	45.4	302	2	Q4K6B2_PSEFP
31	44	45.4	313	2	P90433_SIVCZ

32	44	45.4	332	2	Q4NFI4_9MICC
33	44	45.4	347	2	Q8G7M6_BIFLO
34	44	45.4	371	2	Q53MP5_ORYSA
35	44	45.4	377	2	Q82PX5_STRAW
36	44	45.4	472	2	Q4PED7_USTMA
37	44	45.4	500	2	Q6AKH3_DESPS
38	44	45.4	648	2	Q9P888_GIBFU
39	44	45.4	764	2	Q4WK00_ASPFU
40	44	45.4	1017	1	Q6VG40_SIVCZ
41	44	45.4	1019	1	POL_SIVS4
42	44	45.4	1019	2	P89154_SIVCZ
43	44	45.4	1019	2	Q7ZBR7_SIVCZ
44	44	45.4	1019	2	Q7ZBR5_SIVCZ
45	44	45.4	1058	1	POL_HVZD2
46	44	45.4	1387	2	Q5BFN8_EMENI
47	44	45.4	1928	2	Q8D674_VIBU
48	43.5	44.8	561	2	Q6MDL6_PARUV
49	43.5	44.8	657	2	Q8L07_PSRPK
50	43	44.3	63	2	Q98183_MCVI
51	43	44.3	113	2	Q8VLJ9_9BACT
52	43	44.3	124	2	Q48532_LEPDI
53	43	44.3	215	2	Q6AM22_DESPS
54	43	44.3	216	1	PSAL_SPIOL
55	43	44.3	217	2	Q87115_SIVCZ
56	43	44.3	219	2	Q54C69_DICDI
57	43	44.3	228	2	Q60T23_CABER
58	43	44.3	233	2	Q9KC77_BACHD
59	43	44.3	237	1	PYRF_GLOVI
60	43	44.3	238	2	Q8XSK8_RALSO
61	43	44.3	241	1	YM08_PARTE
62	43	44.3	264	1	Q4LIV5_9ENTR
63	43	44.3	269	1	RIBF_MYCPN
64	43	44.3	304	2	Q5LT69_SILPO
65	43	44.3	340	2	Q8UN04_SIVCZ
66	43	44.3	340	2	Q8UN03_SIVCZ
67	43	44.3	348	2	Q51LJ5_MAGGR
68	43	44.3	349	2	Q7SKK9_9HIV2
69	43	44.3	349	2	Q7SKK8_9HIV2
70	43	44.3	352	1	IDI2_PYRAE
71	43	44.3	352	2	Q830B7_ENTFA
72	43	44.3	369	2	Q90PUL_SIVCZ
73	43	44.3	396	2	Q9XH69_PRUAR
74	43	44.3	454	2	Q897U6_CLOTE
75	43	44.3	472	2	Q89WH9_9HIV2
76	43	44.3	527	1	TOP68_PYRAE
77	43	44.3	536	2	Q8VVM2_ARATH
78	43	44.3	537	2	Q946D4_ARATH
79	43	44.3	609	2	Q856X8_9CAUD
80	43	44.3	900	2	Q8GHS7_PSERE
81	43	44.3	986	2	Q57059_SIVCZ
82	43	44.3	1005	2	Q6Y8X5_9HIV1
83	43	44.3	1022	1	POL_SIVSP
84	43	44.3	1022	2	Q90317_SIVCZ
85	43	44.3	1022	2	Q87956_SIVCZ
86	43	44.3	1022	2	Q87965_SIVCZ
87	43	44.3	1022	2	Q89620_SIVCZ
88	43	44.3	1022	2	Q88135_SIVCZ
89	43	44.3	1023	2	Q5AYH8_EMENI
90	43	44.3	1056	2	Q04097_SIVCZ
91	43	44.3	1057	1	POL_SIVAI
92	43	44.3	1150	2	P90246_9RETR
93	43	44.3	1150	2	Q64F60_9RETR
94	43	44.3	1157	2	Q527F0_MAGGR
95	43	44.3	1226	2	Q6H0K6_9HIV2
96	43	44.3	1583	1	MIS4_SCHPO
97	43	44.3	1896	1	VITI_PEPAM
98	42.5	43.8	88	2	Q6YS15_ORYSA
99	42.5	43.8	275	2	Q8ZFX2_YERPE
100	42.5	43.8	275	2	Q66C47_YERPS
101	42.5	43.8	303	2	Q4JCG1_SULAC
102	42.5	43.8	307	2	Q76IR3_SULAC
103	42.5	43.8	595	2	Q6N169_CORDI
104	42.5	43.8	2281	2	Q6RKL2_GIBMO

Q4NFI4	arthrobacte
Q8G7M6	bifidobacte
Q53MP5	oryza sativ
Q82PX5	streptomyce
Q4PED7	ustilago ma
Q6AKH3	desulfotale
Q9P888	gibberella
Q4WK00	aspergillus
Q6VG40	chimpanzee
P12502	simian immu
P89154	chimpanzee
Q7ZBR7	chimpanzee
Q7ZBR5	chimpanzee
P15833	human immun
Q5BFN8	aspergillus
Q8D674	vibrio vuln
Q6MDL6	parachlamyd
Q88107	pseudomonas
Q98183	molluscum c
Q8VLJ9	uncultured
Q48532	leptothrix
Q6AM22	desulfotale
Q41385	spinacia ol
Q87115	chimpanzee
Q54669	dictyosteli
Q60T23	caenorhabdi
Q9KC77	bacillus ha
Q7nk22	gloeobacter
Q8XSK8	raletonia s
P15609	paramecium
Q4LIV5	salmonella
P75587	mycoplasma
Q5LT69	silicibacte
Q8UN04	chimpanzee
Q8UN03	chimpanzee
Q51LJ5	magnaporthe
Q7SKK9	human immun
Q7SKK8	human immun
Q8zyf6	pyrobaculum
Q830B7	enterococcu
Q90PUL	chimpanzee
Q9XH69	prunus arme
Q897U6	clostridium
Q89WH9	human immun
Q8ZVM0	pyrobaculum
Q8VVM2	arabidopsis
Q946D4	arabidopsis
Q856X8	pycobacteri
Q8GHS7	pseudomonas
Q57059	chimpanzee
Q6Y8X5	human immun
P19505	simian immu
Q90317	chimpanzee
Q87956	chimpanzee
Q87965	chimpanzee
Q89620	chimpanzee
Q88135	chimpanzee
Q5AYH8	aspergillus
Q04097	chimpanzee
Q02836	simian immu
P90246	feline immu
Q64F60	feline immu
Q527F0	magnaporthe
Q6H0K6	human immun
Q99725	schizosacch
Q98M0	periplaneta
Q6YS15	oryza sativ
Q8ZFX2	versinia pe
Q66C47	versinia ps
Q4JCG1	sulfolobus
Q76IR3	sulfolobus
Q6N169	corynebacte
Q6RKL2	gibberella

105	42	43.3	98	2	Q5P854_AZOSE	Q5P854_azoarcus sp	178	41	42.3	251	2	Q5HLW9_STAEO	Q5hlw9 staphylococ
106	42	43.3	114	2	Q58304_PRRHO	Q58304_pyrococcus	179	41	42.3	251	2	Q8CRE9_STAEP	Q8cre9 staphylococ
107	42	43.3	173	2	Q8D9X7_VIBVU	Q8d9x7_vibrio vuln	180	41	42.3	256	2	Q6GEF3_STAAR	Q6gef3 staphylococ
108	42	43.3	174	2	Q4VP81_9GAMM	Q4vp81_wolbachia p	181	41	42.3	257	1	TRMD_AQUAE	Q67463 aquifex ae
109	42	43.3	189	2	Q9MFE1_BETVU	Q9mfe1_beta vulgar	182	41	42.3	264	2	Q8YUM2_ANASP	Q8yum2 anabaena sp
110	42	43.3	226	2	Q8KPF28_CHLUE	Q8kpf28_chlorobium	183	41	42.3	266	2	Q53002_ECOLI	Q53002 escherichia
111	42	43.3	238	2	Q835J7_ENTFPA	Q835j7_enterococu	184	41	42.3	266	2	Q65FV4_BACCL	Q65fv4 bacillus li
112	42	43.3	244	2	Q8PPV5_XANAC	Q8ppv5_xanthomonas	185	41	42.3	271	2	Q4MVQ5_BACCE	Q4mvq5 bacillus ce
113	42	43.3	266	2	Q81D70_BACCR	Q81d70_bacillus ce	186	41	42.3	272	2	Q72XP5_BACCI	Q72xp5 bacillus ce
114	42	43.3	296	2	Q7V459_PROCHLORO	Q7v459_prochloroco	187	41	42.3	278	2	Q9ADV0_9SPHN	Q9adv0 erythroba
115	42	43.3	315	2	Q6HI35_BACHK	Q6hi35_bacillus th	188	41	42.3	288	2	Q7V391_PROMP	Q7v391 prochloroco
116	42	43.3	361	2	Q9RKM5_STRCO	Q9rkm5_streptomyce	189	41	42.3	293	1	ARGB_SYNXP	Q7u5c7 synechococ
117	42	43.3	361	2	Q4WGN3_ASPFU	Q4wgn3_aspergillus	190	41	42.3	305	2	Q6A7C7_PROAC	Q6a7c7 propionibac
118	42	43.3	364	2	Q5E0T4_VIBF1	Q5e0t4_vibrio fisc	191	41	42.3	306	2	Q7D906_MYCTU	Q7d906 mycobacteri
119	42	43.3	379	2	Q8PRC6_XANAC	Q8prc6_xanthomonas	192	41	42.3	306	2	Q05576_MYCTU	Q05576 mycobacteri
120	42	43.3	386	1	ETR1_CANTR	Q8wzm3 candida tro	193	41	42.3	306	2	Q7U0W3_MYCBO	Q7u0w3 mycobacteri
121	42	43.3	386	1	ETR2_CANTR	Q6uuh2 oryza sativ	194	41	42.3	325	2	Q4ZSC3_PSESY	Q4zsc3 pseudomonas
122	42	43.3	391	1	Q6U0H2_ORYSA	Q6uuh2_oryza sativ	195	41	42.3	328	2	Q6NEV8_CORDI	Q6nev8 corynebacte
123	42	43.3	416	1	HEM1_AERPE	Q9y9j2 aeropyrum p	196	41	42.3	331	2	Q4JU09_CORJK	Q4ju09 corynebacte
124	42	43.3	438	2	Q9MFE2_BETVU	Q9mfe2_beta vulgar	197	41	42.3	333	2	Q7MR85_WOLSU	Q7mrr5 wolinnella s
125	42	43.3	443	2	Q89FY2_BRAJA	Q89fy2_bradyrhizob	198	41	42.3	338	2	Q4LPN5_9BURK	Q4lpn5 burkholderi
126	42	43.3	454	2	Q35213_OENBE	Q35213_oenothera b	199	41	42.3	345	2	Q5XHL5_XENLA	Q5xhl5 xenopus lae
127	42	43.3	462	2	Q4LYF0_9BURK	Q4lyf0_burkholderi	200	41	42.3	369	2	Q5NTF8_9BACT	Q5ntf8 uncultured
128	42	43.3	500	2	Q61FS2_CABER	Q61fs2_caenorhabdi	201	41	42.3	375	2	Q6QW16_AZOBR	Q6qw16 azospirilli
129	42	43.3	519	2	Q9SRY9_ARATH	Q9sry9_arabidopsis	202	41	42.3	397	1	TRPB1_ARCFU	Q28672 archaeglob
130	42	43.3	580	2	Q89RH2_BRAJA	Q89rh2_bradyrhizob	203	41	42.3	403	2	Q88NU2_PSEPK	Q88nu2 pseudomonas
131	42	43.3	581	2	Q43856_VICFA	Q43856_vicia faba	204	41	42.3	419	1	CBPAL_HUMAN	Q15085 homo sapien
132	42	43.3	612	2	Q830J7_ENTFPA	Q830j7_enterococu	205	41	42.3	419	2	Q53XU0_HUMAN	Q53xu0 homo sapien
133	42	43.3	628	1	SYF_SYNY3	Q73942 synechocyst	206	41	42.3	440	2	Q5XLI3_AERHY	Q5xli3 aeromonas h
134	42	43.3	637	2	Q6BET2_CAEEL	Q6bet2_caenorhabdi	207	41	42.3	450	2	Q91938_XENLA	Q91938 xenopus lae
135	42	43.3	635	2	Q19186_CAEEL	Q19186_caenorhabdi	208	41	42.3	450	2	Q6DFP3_XENLA	Q6dff3 xenopus lae
136	42	43.3	658	2	Q91AC1_BRARE	Q91ac1_brachydanio	209	41	42.3	450	2	Q6P2X7_XENTR	Q6p2x7 xenopus tro
137	42	43.3	711	2	Q4UPE6_XANCP	Q4upe6_xanthomonas	210	41	42.3	462	2	Q8TMMO_METAC	Q8tmm0_methanosarc
138	42	43.3	711	2	Q8P3W9_XANCP	Q8p3w9_xanthomonas	211	41	42.3	469	2	Q8MMK2_DUGJA	Q8mmk2 dugesia jap
139	42	43.3	711	2	Q8PFF2_XANAC	Q8pff2_xanthomonas	212	41	42.3	472	2	Q7U0G1_MYCBO	Q7u0g1 mycobacteri
140	42	43.3	716	2	Q6CXN1_KLULA	Q6cxn1_kluyveromyc	213	41	42.3	472	2	Q50438_MYCTU	Q50438 mycoba
141	42	43.3	746	2	Q4LINE9_9BURK	Q4line9_burkholderi	214	41	42.3	474	2	Q754V5_ASHGO	Q754v5 ashbya gos
142	42	43.3	776	2	Q6K683_ORYSA	Q6k683_oryza sativ	215	41	42.3	476	2	Q4WL94_ASPFU	Q4wl94 aspergillus
143	42	43.3	805	2	Q91538_XANOR	Q91538_xanthomonas	216	41	42.3	491	2	Q5L085_GSKA	Q5l085 geobacillus
144	42	43.3	822	2	Q33831_THEMA	Q33831_thermotoga	217	41	42.3	496	2	Q6CER7_YARLI	Q6cer7 yarrowia li
145	42	43.3	848	2	Q4RVV3_TETNG	Q4rvv3_tetradodon n	218	41	42.3	503	2	Q4WDM9_ASPFU	Q4wdm9 aspergillus
146	42	43.3	863	2	Q9ST50_MAIZE	Q9st50_zea mays (m	219	41	42.3	519	2	Q4WM15_ASPFU	Q4wm15 aspergillus
147	42	43.3	1015	2	Q87AHL1_SIVCZ	Q87ahl1_chimpanzee	220	41	42.3	552	2	Q83932_9RETR	Q83932 ovine lenti
148	42	43.3	1047	1	POL_SIVAI	P27373 simian immu	221	41	42.3	552	2	Q83933_9RETR	Q83933 ovine lenti
149	42	43.3	1049	2	Q9XBP6_MYXXA	Q9xbp6_myxococcus	222	41	42.3	552	2	Q83934_9RETR	Q83934 ovine lenti
150	42	43.3	1061	1	POL_SIVAT	P05895 simian immu	223	41	42.3	552	2	Q90892_CHICK	Q90892 gallus gall
151	42	43.3	1241	1	Q81TW1_BACAN	Q81tw1_bacillus an	224	41	42.3	576	2	Q90892_CHICK	Q90892 gallus gall
152	42	43.3	1370	2	Q4W7J1_TOBAC	Q4w7j1_nicotiana t	225	41	42.3	594	2	Q4IH39_GIBZE	Q4ih39 gluconobact
153	42	43.3	1407	2	Q4T188_TETNG	Q4t188_tetradodon n	226	41	42.3	594	2	Q5PRH3_GLUOX	Q5fhr3 gluconobact
154	42	43.3	1471	2	Q54173_DICDI	Q54173_dictyosteli	227	41	42.3	607	2	Q87HD5_VIBPA	Q87hd5 vibrio para
155	42	43.3	2108	2	Q7VEZ3_MYCBO	Q7vez3_mycobacteri	228	41	42.3	622	1	MTLL1_LACLA	P35516 lactococcus
156	42	43.3	2108	2	Q53901_MYCTU	Q53901_mycobacteri	229	41	42.3	664	2	Q6LKN3_PROPR	Q6lkn3 photobacter
157	42	43.3	2282	2	Q8K4S1_MOUSE	Q8k4s1_mus musculu	230	41	42.3	679	2	Q5IU94_MAGGR	Q5iu94 magnaporthe
158	41.5	42.8	168	2	Q9V492_DROME	Q9v492_drosophila	231	41	42.3	695	2	Q4RH59_TETNG	Q4rh59 tetradodon n
159	41.5	42.8	308	1	XERC_CORGL	Q8nnz9 corynebacte	232	41	42.3	715	2	Q6PAA9_XENLA	Q6paa9 xenopus lae
160	41.5	42.8	406	2	Q65EX2_BACLD	Q65ex2_bacillus li	233	41	42.3	721	2	Q8K6D6_STRP3	Q8k6d6 streptococ
161	41.5	42.8	877	2	O15644_TOXGO	O15644_toxoplasma	234	41	42.3	721	2	Q5XAM7_STRP6	Q5xam7 streptococ
162	41	42.3	53	2	Q5WYJ5_NOCFU	Q5wyj5_nocardia fa	235	41	42.3	721	2	Q99YL1_STRPY	Q99yl1 streptococ
163	41	42.3	61	2	Q8VJ58_MYCTU	Q8vj58_mycobacteri	236	41	42.3	721	2	Q7CMY6_STRP8	Q7cm6 streptococ
164	41	42.3	70	1	RK28_CYPAP	P48129 cyanophora	237	41	42.3	801	2	Q5AS91_EMENI	Q5aa91 aspergillus
165	41	42.3	116	2	Q623C8_ORYSA	Q623c8_oryza sativ	238	41	42.3	810	2	Q5TFJ4_HUMAN	Q5tfj4 homo sapien
166	41	42.3	135	2	Q6FSL6_CANGA	Q6fsl6_candida gla	239	41	42.3	889	2	Q9F179_STRGR	Q9f179 streptomyce
167	41	42.3	153	2	Q7CZF6_AGRF5	Q7czf6_agrobacteri	240	41	42.3	897	2	Q8EII3_SHEON	Q8eii3 shewanella
168	41	42.3	173	2	Q5WHI4_BAGSK	Q5whi4_bacillus cl	241	41	42.3	967	2	Q4FP31_9RICK	Q4fp31 candidatus
169	41	42.3	178	2	Q8UPP3_AGRF5	Q8upp3_agrobacteri	242	41	42.3	1007	2	Q699W7_SIVCZ	Q699w7 chimpanzee
170	41	42.3	188	2	Q9PHI6_XYLFA	Q9phi6_xylolla fas	243	41	42.3	1007	2	Q699V9_SIVCZ	Q699v9 chimpanzee
171	41	42.3	192	2	Q8SVU0_ENCCU	Q8svu0_encephalito	244	41	42.3	1018	2	Q91BL9_9HIV2	Q91bl9 human immu
172	41	42.3	203	2	P87305_SCHPO	P87305_schizosacch	245	41	42.3	1039	2	Q87108_SIVCZ	Q87108 chimpanzee
173	41	42.3	237	1	PYRF_ENTFA	Q834e3 enterococu	246	41	42.3	1046	1	POL_SIVAG	P27980 simian immu
174	41	42.3	238	2	P74728_SYNY3	P74728_synechocyst	247	41	42.3	1055	2	Q89928_9HIV2	Q89928 human immu
175	41	42.3	241	2	Q6P8B0_ACTAD	Q6f8b0_acinetobact	248	41	42.3	1060	2	Q5QGQ1_SIVCZ	Q5qgq1 chimpanzee
176	41	42.3	245	2	Q66272_9SPHN	Q66272_erythroba	249	41	42.3	1060	2	Q5QFT5_SIVCZ	Q5qft5 chimpanzee
177	41	42.3	249	2	Q82989_9SPHN	Q82989_erythroba	250	41	42.3	1060	2	Q5QFR3_SIVCZ	Q5qfr3 chimpanzee

251	41	42.3	1060	2	Q5QFL2_SIVCZ	Q5qf12 chimpanzee	324	40	41.2	349	2	Q9RQL6_STRPN	Q9rql6 streptococ
252	41	42.3	1060	2	Q5QFJ6_SIVCZ	Q5qf16 chimpanzee	325	40	41.2	349	2	Q4JYG2_STRPN	Q4jyg2 streptococ
253	41	42.3	1061	2	Q9W699_FUGRU	Q9w699 fugu rubrip	326	40	41.2	349	2	Q4JYN4_STRPN	Q4jyn4 streptococ
254	41	42.3	1061	2	Q76634_9HIV2	Q76634 human immun	327	40	41.2	349	2	Q4JYW5_STRPN	Q4jyw5 streptococ
255	41	42.3	1064	2	Q8UTV7_9HIV2	Q8utv7 human immun	328	40	41.2	349	2	Q4JYV6_STRPN	Q4jyv6 streptococ
256	41	42.3	1123	2	Q9SE99_ARATH	Q9se99 arabidopsis	329	40	41.2	349	2	Q4K034_STRPN	Q4k034 streptococ
257	41	42.3	1123	2	Q9SPU7_ARATH	Q9spu7 arabidopsis	330	40	41.2	349	2	Q4K0A7_STRPN	Q4k0a7 streptococ
258	41	42.3	1124	1	POL_FIVSD	P19028 f pol polyp	331	40	41.2	349	2	Q4K0Q5_STRPN	Q4k0q5 streptococ
259	41	42.3	1140	1	Y376_MYCPN	P75405 mycoplasma	332	40	41.2	349	2	Q4K121_STRPN	Q4k121 streptococ
260	41	42.3	1191	2	Q52484_SERMA	O52484 serratia ma	333	40	41.2	349	2	Q4K121_STRPN	Q4k121 streptococ
261	41	42.3	1820	1	Q54F55_DICDI	O54f55 dictyosteli	334	40	41.2	349	2	Q4K2T4_STRPN	Q4k2t4 streptococ
262	41	42.3	3433	1	PCX_DROME	P18490 drosophila	335	40	41.2	349	2	Q4K2T4_STRPN	Q4k2t4 streptococ
263	40.5	41.8	324	2	Q7T9Q0_GVAV	Q7t9q0 adoxophyes	336	40	41.2	349	2	O54546_STRPN	O54546 streptococ
264	40.5	41.8	504	2	Q6EMC3_ECOLI	Q6emc3 escherichia	337	40	41.2	349	2	Q8DR64_STRPN	Q8dr64 streptococ
265	40.5	41.8	504	2	Q9WTH8_92ZZZ	Q9wth8 plasmid r10	338	40	41.2	349	2	Q7SKK6_9HIV2	Q7skk6 human immun
266	40.5	41.8	507	1	TRAC_PSEAE	P0a134 pseudomonas	339	40	41.2	349	2	O7SKK4_9HIV2	O7skk4 human immun
267	40.5	41.8	507	1	TRA6_SHIFL	P0a135 shigella fl	340	40	41.2	349	2	O7SKJ5_9HIV2	O7skj5 human immun
268	40.5	41.8	507	1	Q7BT57_ECOLI	Q7bt57 escherichia	341	40	41.2	349	2	O7SKJ4_9HIV2	O7skj4 human immun
269	40.5	41.8	507	2	Q79PC0_KLEPN	Q79pc0 klebsiella	342	40	41.2	349	2	O7SKJ3_9HIV2	O7skj3 human immun
270	40.5	41.8	507	2	Q79CH1_PSEAE	Q79ch1 pseudomonas	343	40	41.2	349	2	O7SKJ2_9HIV2	O7skj2 human immun
271	40.5	41.8	889	2	Q4IUH0_AZOV1	Q4iuh0 azotobacter	344	40	41.2	349	2	O7SKJ1_9HIV2	O7skj1 human immun
272	40	41.2	81	2	Q9NDL7_HYDMA	Q9ndl7 hydra magni	345	40	41.2	349	2	O7SKJ0_9HIV2	O7skj0 human immun
273	40	41.2	100	2	Q6VFT3_POXV	Q6vft3 squirrel po	346	40	41.2	349	2	O7SKI8_9HIV2	O7ski8 human immun
274	40	41.2	108	2	Q51C64_ENTHI	O51c64 entamoeba h	347	40	41.2	349	2	O7SKI7_9HIV2	O7ski7 human immun
275	40	41.2	115	2	Q4PN77_IXOSC	Q4pn77 ixodes scap	348	40	41.2	349	2	O7SKI6_9HIV2	O7ski6 human immun
276	40	41.2	129	1	RS9_WOFSU	O7m7r1 wolinnella s	349	40	41.2	349	2	O7SKI5_9HIV2	O7ski5 human immun
277	40	41.2	129	2	Q8DEX7_SYNEL	O8dex7 synchococ	350	40	41.2	349	2	O7SKI4_9HIV2	O7ski4 human immun
278	40	41.2	133	2	Q4ZPV7_PSEY1	Q4zpv7 pseudomonas	351	40	41.2	349	2	O7SKI3_9HIV2	O7ski3 human immun
279	40	41.2	134	2	Q88615_PSESM	Q88615 pseudomonas	352	40	41.2	349	2	O7SKI2_9HIV2	O7ski2 human immun
280	40	41.2	139	2	Q9GSF9_HYDAT	Q9gsf9 hydra atten	353	40	41.2	349	2	O7SKI1_9HIV2	O7ski1 human immun
281	40	41.2	146	2	Q6WMX3_ANAPH	Q6wmx3 anaplasm p	354	40	41.2	351	2	O9E3T2_9HIV2	O9e3t2 human immun
282	40	41.2	154	2	Q9WI01_9HIV2	Q9wi01 human immun	355	40	41.2	351	2	O9E3T4_9HIV2	O9e3t4 human immun
283	40	41.2	158	2	Q9X0A5_THEMA	Q9x0a5 thermotoga	356	40	41.2	351	2	O9E3T4_9HIV2	O9e3t4 human immun
284	40	41.2	165	2	Q5PX28_9CHLO	Q5px28 halimeda tu	357	40	41.2	355	1	TSN10_HUMAN	O8hkn4 uncultured
285	40	41.2	172	2	Q5RI40_BRARE	O5ri40 brachydanio	358	40	41.2	361	2	O8NKN4_9CREN	O8nkn4 uncultured
286	40	41.2	172	2	Q9G8T3_9CRYP	Q9g8t3 rhodomonas	359	40	41.2	368	2	O53605_STAAU	O53605 staphylococ
287	40	41.2	182	2	Q9PSV8_NEUCR	Q9psv8 neurospora	360	40	41.2	368	2	O69LD7_ORYXA	O69ld7 oryza sativ
288	40	41.2	185	2	Q7RVZ3_NEUCR	Q7rvz3 neurospora	361	40	41.2	369	2	O6AGP0_LEIXX	O6agp0 leifsonia x
289	40	41.2	193	1	FTPM1_MOUSE	Q66gt5 mus musculus	362	40	41.2	374	2	O6AGP0_LEIXX	O6agp0 leifsonia x
290	40	41.2	207	1	GST3_CAEEL	O16116 caenorhabdi	363	40	41.2	391	2	O6SEAF_9CAUD	O6seaf lactobacill
291	40	41.2	211	2	P88796_9HIV1	P88796 human immun	364	40	41.2	391	2	O6SEAF_9CAUD	O6seaf lactobacill
292	40	41.2	212	2	Q4KJ2_STRPN	Q4kj2 streptococ	365	40	41.2	394	2	O56919_9HIV1	O56919 human immun
293	40	41.2	216	2	Q9XVJ3_CAEEL	Q9xvj3 caenorhabdi	366	40	41.2	394	2	O56920_9HIV1	O56920 human immun
294	40	41.2	242	1	Y1085_WCTBU	P67157 mycobacteri	367	40	41.2	394	2	O56917_9HIV1	P36875 pisum sativ
295	40	41.2	242	1	Y1114_MYCBO	P67158 mycobacteri	368	40	41.2	395	1	2AAA_PEA	O7zmy2 human immun
296	40	41.2	242	1	Y1114_MYCBO	P67158 mycobacteri	369	40	41.2	395	1	2AAA_PEA	O7zmy2 human immun
297	40	41.2	244	2	Q6ZN98_HUMAN	Q6zn98 homo sapien	370	40	41.2	399	2	O8ID4_FUSNN	O8id4 fusobacteri
298	40	41.2	249	2	Q9A535_CAUCR	Q9a535 caulobacter	371	40	41.2	401	2	O7VVG2_BORPE	O7vvg2 bordetella
299	40	41.2	251	2	Q4KAL0_PSEFS	Q4kal0 pseudomonas	372	40	41.2	401	2	O7WB96_BORPA	O7wb96 bordetella
300	40	41.2	261	2	Q5IR73_9HIV1	Q5ir73 human immun	373	40	41.2	401	2	O7WMR4_BORBR	O7wmr4 bordetella
301	40	41.2	261	2	Q5IR73_9HIV1	Q5ir73 human immun	374	40	41.2	403	2	O8KDE7_CHLTE	O8kde7 chlorobium
302	40	41.2	262	2	Q9HWM4_PSEAE	Q9hwm4 pseudomonas	375	40	41.2	412	2	O8IN73_DROME	O8in73 drosophila
303	40	41.2	265	2	O33643_SALGL	O33643 salmonella	376	40	41.2	419	1	CBPAL_MOUSE	O7tpa8 mus musculu
304	40	41.2	283	2	O6ANH6_DESPS	Q6anh6 desulfotale	377	40	41.2	419	1	CBPAL_MOUSE	O7tpa8 mus musculu
305	40	41.2	288	2	Q7U3V2_SYNPX	Q7u3v2 synchococ	378	40	41.2	419	1	CBPAL_MOUSE	O7tpa8 mus musculu
306	40	41.2	297	2	Q7UQ24_RHOBA	Q7uq24 rhodopirell	379	40	41.2	428	2	O9D4K7_MOUSE	O9d4k7 mus musculu
307	40	41.2	306	2	Q9Z5G1_MYCLE	Q9z5g1 mycobacteri	380	40	41.2	428	2	O9D4K7_MOUSE	O9d4k7 mus musculu
308	40	41.2	310	2	Q7SKI9_9HIV2	Q7ski9 human immun	381	40	41.2	429	2	O38614_BPTU2	O38614 lactococcus
309	40	41.2	312	2	Q978P0_THEVO	Q978p0 thermoplasma	382	40	41.2	429	2	O38614_BPTU2	O38614 lactococcus
310	40	41.2	325	2	O5QY78_DILO	O5qyt8 idiomarina	383	40	41.2	429	2	MTAL_ACEPA	O52702 acetobacter
311	40	41.2	330	2	Q84M50_ORYSA	Q84m50 oryza sativ	384	40	41.2	429	1	MTAL_ACEPA	O52702 acetobacter
312	40	41.2	336	2	Q9WYF8_THEMA	Q9wyf8 thermotoga	385	40	41.2	429	1	MTAL_ACEPA	O52702 acetobacter
313	40	41.2	344	2	Q9HN36_HALISA	Q9hn36 halobacteri	386	40	41.2	429	2	O37969_9CAUD	O37969 bacterioph
314	40	41.2	348	2	O6L5P7_STRMT	Q6l5p7 streptococ	387	40	41.2	429	2	O8LTP4_9CAUD	O8ltp4 lactococcus
315	40	41.2	348	2	O6L5R8_STROR	Q6l5r8 streptococ	388	40	41.2	429	2	O9FZS8_9CAUD	O9fzs8 lactococcus
316	40	41.2	348	2	Q83Y74_STRGN	Q83yt4 streptococ	389	40	41.2	429	2	O9FZS8_9CAUD	O9fzs8 lactococcus
317	40	41.2	348	2	O9Z641_STRPN	Q9z641 streptococ	390	40	41.2	429	2	O9T1Z3_9CAUD	O9t1z3 bacterioph
318	40	41.2	348	2	Q7SS3_VIBPA	Q7ss3 vibrio para	391	40	41.2	429	2	O98700_SIVCZ	O98700 chimpanzee
319	40	41.2	349	2	O50Z69_ENTHI	O50z69 entamoeba h	392	40	41.2	439	2	O7R3Z8_GIALA	O7r3z8 giardia lam
320	40	41.2	349	2	O54611_STRPN	O54611 streptococ	393	40	41.2	443	2	O7ULE0_RHOBA	O7ule0 rhodopirell
321	40	41.2	349	2	P97006_STRPN	P97006 streptococ	394	40	41.2	457	2	O7XBC8_MAIZE	O7xbc8 sea mays (m
322	40	41.2	349	2	O84A36_STRPN	O84a36 streptococ	395	40	41.2	469	2	O8AAL8_BACTN	O8aal8 bacteroides
323	40	41.2	349	2	Q9AH82_STRPN	Q9ah82 streptococ	396	40	41.2	476	2	O9WI00_9HIV2	O9wi00 human immun
					Q9AH82_STRPN	Q9ah82 streptococ				488	2	Q93383_CAEEL	Q93383 caenorhabdi
					Q9AH82_STRPN	Q9ah82 streptococ				489	2	O62100_CAEEL	O62100 caenorhabdi
					Q9AH82_STRPN	Q9ah82 streptococ				495	2	O4WR57_ASPPU	O4wr57 aspergillus
					Q9AH82_STRPN	Q9ah82 streptococ				497	2	O5CSN2_CRYPV	O5csn2 cryptocsori
					Q9AH82_STRPN	Q9ah82 streptococ				497	2	O5CJ09_CRYHO	O5cj09 cryptocsori

397	40	41.2	500	2	Q4V7B5_RAT	Q4v7b5	rattus norv	470	41.2	1060	2	Q90EX8_SIVCZ	Q90ex8	simian immu
398	40	41.2	520	2	Q4IE36_GIBZE	Q4ie36	gibberella	471	41.2	1060	2	Q8USP8_9PLVG	Q8usp8	simian-huma
399	40	41.2	521	2	Q82507_ARATH	Q82507	arabidopsis	472	41.2	1060	2	Q07387_SIVCZ	Q07387	chimpanzee
400	40	41.2	525	2	Q6CCQ9_YARLI	Q6ccq9	yarrowia li	473	41.2	1060	2	Q5W9H4_SIVCZ	Q5w9h4	chimpanzee
401	40	41.2	527	2	Q25564_HELPY	Q25564	helicobacte	474	41.2	1060	2	Q5QGP0_SIVCZ	Q5qgp0	chimpanzee
402	40	41.2	535	2	Q5NE19_LYCES	Q5ne19	lycopersico	475	41.2	1060	2	Q5QGH9_SIVCZ	Q5qgh9	chimpanzee
403	40	41.2	540	2	Q6ALB3_DESPS	Q6alb3	desulfotale	476	41.2	1060	2	Q5QGH0_SIVCZ	Q5qgh0	chimpanzee
404	40	41.2	543	2	Q94185_ASPOR	Q94185	aspergillus	477	41.2	1060	2	Q5QGF7_SIVCZ	Q5qgf7	chimpanzee
405	40	41.2	553	2	Q7MGH7_BORBR	Q7mgh7	bordetella	478	41.2	1060	2	Q5QGB6_SIVCZ	Q5qgb6	chimpanzee
406	40	41.2	553	2	Q7W512_BORPA	Q7w512	bordetella	479	41.2	1060	2	Q5QGA9_SIVCZ	Q5qga9	chimpanzee
407	40	41.2	553	2	Q7VVX5_BORPE	Q7vvx5	bordeella	480	41.2	1060	2	Q5QGA9_SIVCZ	Q5qga9	chimpanzee
408	40	41.2	556	2	Q6ONX9_CAEHR	Q6onx9	caenorhabdi	481	41.2	1060	2	Q5QGA9_SIVCZ	Q5qga9	chimpanzee
409	40	41.2	560	2	Q06347_9HIV2	Q06347	human immu	482	41.2	1060	2	Q5QGA9_SIVCZ	Q5qga9	chimpanzee
410	40	41.2	565	2	Q4IRR7_GIBZE	Q4irr7	gibberella	483	41.2	1060	2	Q5QGF7_SIVCZ	Q5qgf7	chimpanzee
411	40	41.2	583	2	Q7YTY5_DROME	Q7yty5	drosophila	484	41.2	1060	2	Q5QFV2_SIVCZ	Q5qfv2	chimpanzee
412	40	41.2	584	2	Q5YXR0_9TRYP	Q5yxr0	trypanosoma	485	41.2	1060	2	Q5QFV2_SIVCZ	Q5qfv2	chimpanzee
413	40	41.2	586	2	Q828W4_STEAM	Q828w4	streptomyce	486	41.2	1060	2	Q5QFS2_SIVCZ	Q5qfs2	chimpanzee
414	40	41.2	601	2	Q8MS83_DROME	Q8mas8	drosophila	487	41.2	1060	2	Q5QF04_SIVCZ	Q5qf04	chimpanzee
415	40	41.2	601	2	Q9W3R0_DROME	Q9w3r0	drosophila	488	41.2	1060	2	Q5QF04_SIVCZ	Q5qf04	chimpanzee
416	40	41.2	612	1	AMYG_ASPOR	P36914	aspergillus	489	41.2	1060	2	Q5QFM3_SIVCZ	Q5qfm3	chimpanzee
417	40	41.2	622	2	Q6BML3_DEBHA	Q6bml3	debariomyce	490	41.2	1060	2	Q5QFM3_SIVCZ	Q5qfm3	chimpanzee
418	40	41.2	623	1	ARE1_SACBA	Q87613	saccharomyc	491	41.2	1060	2	Q5QFK8_SIVCZ	Q5qfk8	chimpanzee
419	40	41.2	666	2	Q72V70_BACC1	Q72v70	bacillus ce	492	41.2	1060	2	Q5QF19_SIVCZ	Q5qf19	chimpanzee
420	40	41.2	666	2	Q81X11_BACAN	Q81x11	bacillus an	493	41.2	1060	2	Q5IZC2_SIVCZ	Q5izc2	chimpanzee
421	40	41.2	666	2	Q6HBP2_BACHK	Q6hbp2	bacillus th	494	41.2	1090	2	Q4XZK5_PLACH	Q4xzk5	plasmodium
422	40	41.2	678	2	Q6GNS6_XENLA	Q6gns6	xenopus lae	495	41.2	1092	2	Q9MB43_ADICA	Q9mb43	adiantum ca
423	40	41.2	688	2	Q9FZR2_9CAUD	Q9fzr2	mycoplasma	496	41.2	1105	2	Q917S4_DROME	Q917s4	drosophila
424	40	41.2	722	2	Q6D2D3_ERWCT	Q6d2d3	erwinia car	497	41.2	1110	2	Q4P945_USTMA	Q4p945	utilago ma
425	40	41.2	753	2	Q4LTC0_9BURK	Q4ltc0	burkholderi	498	41.2	1123	2	Q66933_9RETR	Q66933	feline immu
426	40	41.2	772	1	NUTD3_ARATH	Q81831	arabidopsis	499	41.2	1139	2	Q6C802_YARLI	Q6c802	yarrowia li
427	40	41.2	819	2	Q5HIC9_XANOR	Q5hic9	xanthomonas	500	41.2	1210	2	Q54YU8_DICDI	Q54y8	dictyosteli
428	40	41.2	822	2	Q6HER9_ORYSA	Q6her9	oryza sativ	501	41.2	1485	2	Q69U98_ORYSA	Q69u98	oryza sativ
429	40	41.2	832	2	Q5FAN0_APIME	Q5fan0	apis mellif	502	41.2	1552	2	Q4N585_THEPA	Q4n585	theileria p
430	40	41.2	898	2	Q9W1S3_DROME	Q9wis3	drosophila	503	41.2	1629	2	Q7X7Y2_ORYSA	Q7x7y2	oryza sativ
431	40	41.2	933	2	Q677U1_9VIRU	Q677u1	lymphocysti	504	41.2	1706	2	Q5VZA9_HUMAN	Q5vza9	homo sapien
432	40	41.2	936	2	Q72J29_SIVCZ	Q72j29	simian immu	505	41.2	1712	1	C04A2_HUMAN	P05572	homo sapien
433	40	41.2	941	2	Q5BCP6_EMENI	Q5bcp6	aspergillus	506	41.2	2107	1	Y4J2_CAEEL	P34367	caenorhabdi
434	40	41.2	978	2	Q6G892_STAAS	Q6g892	staphylococ	507	41.2	2609	2	Q74390_SCHPO	Q74390	schizosacch
435	40	41.2	978	2	Q6GFL3_STAAC	Q6gfl3	staphylococ	508	41.2	3434	2	Q5WPU4_9FLAV	Q5wpu4	usutu virus
436	40	41.2	978	2	Q5HET2_STAAC	Q5het2	staphylococ	509	41.2	3434	2	Q5WPU5_9FLAV	Q5wpu5	usutu virus
437	40	41.2	978	2	Q7AOK3_STAAM	Q7aok3	staphylococ	510	41.2	10625	2	Q6W5Q0_9ACTO	Q6w5q0	streptomyce
438	40	41.2	978	2	Q7RAV5_STAAN	Q7rav5	staphylococ	511	39.5	2025	2	Q9FGP7_ARATH	Q9fgp7	arabidopsis
439	40	41.2	978	2	Q99T35_STAAM	Q99t35	staphylococ	512	39.5	202	2	Q5WLD4_BACSK	Q5wld4	bacillus cl
440	40	41.2	1005	2	Q6Y8U8_9HIV1	Q6y8u8	human immu	513	39.5	318	2	Q5BG06_EMENI	Q5bg06	aspergillus
441	40	41.2	1007	2	Q6GQF6_9HIV1	Q6gqf6	human immu	514	39.5	363	2	Q4Q5M2_LBIMA	Q4q5m2	leishmania
442	40	41.2	1009	1	POL_SIVGB	P22382	simian immu	515	39.5	431	2	Q52DS4_MAGGR	Q52ds4	magnaporthe
443	40	41.2	1009	2	Q6Y329_9HIV1	Q6y329	simian immu	516	39.5	475	2	Q5VP06_DROME	Q5vp06	drosophila
444	40	41.2	1012	2	Q6Y8U0_9HIV1	Q6y8u0	human immu	517	39.5	580	2	Q4IDQ2_GIBBEZ	Q4idq2	gibberella
445	40	41.2	1016	2	Q8W0Y3_MAIZE	Q8w0y3	zea mays (m	518	39.5	642	2	Q4ZB95_PSESY	Q4z95	pseudomonas
446	40	41.2	1027	2	Q9WPP5_SIVCZ	Q9wpp5	chimpanzee	519	39.5	642	2	Q9HZ16_PSEAE	Q9hzi6	pseudomonas
447	40	41.2	1033	2	Q79666_9HIV1	Q79666	human immu	520	39.5	718	2	Q606U4_METCA	Q606u4	methylcococ
448	40	41.2	1034	1	POL_HV2CA	P24107	human immu	521	39.5	722	1	FIN1_SCHPO	O13839	schizosacch
449	40	41.2	1035	2	Q73194_9HIV2	Q73194	human immu	522	39.5	964	2	Q6XGE6_ECOLI	Q6xge6	escherichia
450	40	41.2	1035	2	Q6R793_9HIV2	Q6r793	human immu	523	39.5	1746	2	Q22830_CAEEL	Q22830	caenorhabdi
451	40	41.2	1035	2	Q6R784_9HIV2	Q6r784	human immu	524	39.5	4643	2	Q51XM7_MAGGR	Q51xm7	magnaporthe
452	40	41.2	1036	1	POL_HV2RO	P04584	human immu	525	39	46	2	Q74ED1_GROSL	Q74ed1	geobacter s
453	40	41.2	1040	2	Q9T014_ARATH	Q9t014	arabidopsis	526	39	70	2	Q605U7_METCA	Q605u7	methylcococ
454	40	41.2	1043	2	Q90DD2_SIVCZ	Q90dd2	chimpanzee	527	39	81	2	Q9NDL5_9CNID	Q9ndl5	tina formos
455	40	41.2	1049	1	POL_HV2G1	P18042	human immu	528	39	81	2	Q9NDL8_HYDEC	Q9ndl8	hydractinia
456	40	41.2	1054	1	POL_SIVMK	P05897	simian immu	529	39	81	2	Q9NDL9_9CNID	Q9ndl9	eirene sp.
457	40	41.2	1054	2	Q5TYK3_SIVCZ	Q5tyk3	chimpanzee	530	39	98	2	Q67PB8_NYTH	Q67pb8	symbiobacte
458	40	41.2	1055	1	POL_HV2ST	P20876	human immu	531	39	98	2	Q82R21_STAM	Q82r21	streptomyce
459	40	41.2	1055	2	Q9YTU0_9HIV2	Q9ytu0	human immu	532	39	108	2	Q8BSD1_MOUSE	Q8bsd1	mus musculu
460	40	41.2	1056	1	POL_SIVM1	P05896	simian immu	533	39	126	2	Q8N9N4_HUMAN	Q8n9n4	homo sapien
461	40	41.2	1059	2	Q56317_9PLVG	Q56317	simian-huma	534	39	136	2	Q6WQ04_GLOTR	Q6wq04	glossophyllu
462	40	41.2	1059	2	O11404_9PLVG	O11404	simian-huma	535	39	40.2	2	Q5ZFK0_POPTN	Q5zfk0	populus tre
463	40	41.2	1059	2	O56320_9PLVG	O56320	simian-huma	536	39	40.2	2	Q6D8E8_ERMCT	Q6d8e8	erwinia car
464	40	41.2	1059	2	Q76630_9PLVG	Q76630	human immu	537	39	154	2	Q5AXF9_DICDI	Q5axf9	dictyosteli
465	40	41.2	1059	2	O56613_9PLVG	O56613	simian-huma	538	39	165	2	Q5PX26_9CHLO	Q5px26	halimeda tu
466	40	41.2	1059	2	Q5XLN1_9PLVG	Q5xln1	simian-huma	539	39	174	2	Q4GX23_9CHLO	Q4gx23	halimeda hu
467	40	41.2	1060	2	Q90EX3_SIVCZ	Q90ex3	simian immu	540	39	40.2	2	Q96ZS9_SULTO	Q96z89	sulfolobus
468	40	41.2	1060	2	Q87706_SIVCZ	Q87706	chimpanzee	541	39	40.2	2	Q4VP36_FRATT	Q4vp36	franciella
469	40	41.2	1060	2	Q88016_SIVCZ	Q88016	chimpanzee	542	39	40.2	2	Q4VP38_FRATT	Q4vp38	franciella

543	39	40.2	174	2	Q4VP41 FRATU	Q4VP41 francisella	616	39	40.2	294	2	Q94FC9_MAIZE	Q94fc9 zea mays (m
544	39	40.2	174	2	Q4VP45 FRATU	Q4VP45 francisella	617	39	40.2	294	2	Q94FD0_MAIZE	Q94fd0 zea mays (m
545	39	40.2	174	2	Q4VP47 FRATU	Q4VP47 francisella	618	39	40.2	296	2	Q6R4X2_ZEAMP	Q6r4x2 zea mays (s
546	39	40.2	174	2	Q4VP40 FRATU	Q4VP40 francisella	619	39	40.2	296	2	Q6R4X4_ZEAMP	Q6r4x4 zea mays (s
547	39	40.2	183	2	Q9LUES ARATH	Q9lues arabidopsis	620	39	40.2	296	2	Q6R4Y1_ZEAMP	Q6r4y1 zea mays (s
548	39	40.2	188	2	Q5Z7Q8 ORYSA	Q5z7q8 oryza sativ	621	39	40.2	296	2	Q6R4Y3_ZEAMP	Q6r4y3 zea mays (s
549	39	40.2	195	2	Q9S2G3 STROCO	Q9s2g3 streptomyce	622	39	40.2	296	2	Q6R4Y4_ZEALU	Q6r4y4 zea luxuria
550	39	40.2	205	1	COAE GLOVI	Q7nlp9 gloebacter	623	39	40.2	296	2	Q6R4Y5_ZEALU	Q6r4y5 zea luxuria
551	39	40.2	205	2	Q55LP5 CRYNE	O55lp5 cryptococcu	624	39	40.2	296	2	Q6R4Y6_ZEALU	Q6r4y6 zea luxuria
552	39	40.2	206	2	Q5NGD3 FRANTT	O5ngd3 francisella	625	39	40.2	296	2	Q6R4Y7_ZEALU	Q6r4y7 zea luxuria
553	39	40.2	206	2	Q6MEP0 PARUM	Q6mep0 parachlamyd	626	39	40.2	296	2	Q6R4Y8_ZEALU	Q6r4y8 zea luxuria
554	39	40.2	212	2	Q4JU70 CORJK	Q4ju70 corynebacte	627	39	40.2	296	2	Q6R4Y9_ZEALU	Q6r4y9 zea luxuria
555	39	40.2	214	2	Q7X3Y6 VIBUO	Q7x3y6 vibrio vuln	628	39	40.2	296	2	Q6R4Z1_ZEALU	Q6r4z1 zea luxuria
556	39	40.2	214	2	O5LGQ5 BACFN	O5lgq5 bacteroides	629	39	40.2	296	2	Q93W75_MAIZE	Q93w75 zea mays (m
557	39	40.2	214	2	Q6F7K2 ACIAD	O6f7k2 acinetobact	630	39	40.2	296	2	Q93W76_MAIZE	Q93w76 zea mays (m
558	39	40.2	214	2	Q64XK3 BACFR	O64xk3 bacteroides	631	39	40.2	296	2	Q94FC7_MAIZE	Q94fc7 zea mays (m
559	39	40.2	215	2	Q6D962 ERWCT	Q6d962 erwinia car	632	39	40.2	296	2	Q94FC8_MAIZE	Q94fc8 zea mays (m
560	39	40.2	217	2	Q89ZQ0 BACTN	Q89zq0 bacteroides	633	39	40.2	298	2	Q6R4W7_ZEAMP	Q6r4w7 zea mays (s
561	39	40.2	218	2	Q526P1 MAGGR	O526p1 magnaporth	634	39	40.2	298	2	Q6R4W8_ZEAMP	Q6r4w8 zea mays (s
562	39	40.2	220	1	NADD YERPE	Q8zdg1 yersinia pe	635	39	40.2	298	2	Q6R4X0_ZEAMP	Q6r4x0 zea mays (s
563	39	40.2	220	2	Q66D66 YERPS	O66d66 versinia ps	636	39	40.2	298	2	Q6R4X1_ZEAMP	Q6r4x1 zea mays (s
564	39	40.2	227	2	Q64SE4 _SALA	O64se4 oedipina po	637	39	40.2	298	2	Q6R4X7_ZEAMM	Q6r4x7 zea mays (s
565	39	40.2	234	2	Q7N956 PHOLL	Q7n956 photorhabdu	638	39	40.2	298	2	Q6R4X9_ZEAMM	Q6r4x9 zea mays (s
566	39	40.2	238	2	O12158 9HIV1	O12158 human immun	639	39	40.2	298	2	Q6R4Y0_ZEAMM	Q6r4y0 zea mays (s
567	39	40.2	240	1	HIS4 LISMF	Q722y6 listeria mo	640	39	40.2	298	2	Q53QG3_ORYSA	Q53qg3 oryza sativ
568	39	40.2	240	1	UNG_LISMO	Q8y9g4 listeria mo	641	39	40.2	300	2	Q6R4X5_ZEAMM	Q6r4x5 zea mays (s
569	39	40.2	240	1	UNG_MYCPN	Q75536 mycoplasma	642	39	40.2	301	2	Q64WZ2_BACFR	Q64wz2 bacteroides
570	39	40.2	241	2	Q6U023_ORYSA	Q6uu23 oryza sativ	643	39	40.2	301	2	Q82GR2_STRAW	Q82gr2 streptomyce
571	39	40.2	253	2	Q52575_PSEX	O52575 pseudomonas	644	39	40.2	304	2	Q81OP1_MOUSE	Q81op1 mus musculus
572	39	40.2	256	1	VMAT ORSVW	Q84131 ovine respi	645	39	40.2	304	2	Q82845_NITEU	Q82845 nitrosomona
573	39	40.2	256	2	Q7Y1L0 BR5VA	Q7y1l0 bovine resp	646	39	40.2	311	1	LEUO_ECOLI	Q7zyd7 xenila
574	39	40.2	256	2	Q9YS27_OMONO	Q9ys27 bovine resp	647	39	40.2	314	2	Q7AHR5_ECOS7	Q7ahr5 escherichia
575	39	40.2	268	2	Q6TSX5_9POAL	Q6tsx5 bouteloua h	648	39	40.2	314	2	Q89D85_BRAJA	Q89d85 bradyrhizob
576	39	40.2	268	2	Q6TSX6_9POAL	Q6tsx6 bouteloua h	649	39	40.2	314	2	Q8FL73_ECOL6	Q8fl73 escherichia
577	39	40.2	268	2	Q6TSX7_9POAL	Q6tsx7 bouteloua h	650	39	40.2	314	1	NIA_CHLVU	Q81l70 chlorella v
578	39	40.2	268	2	Q6TSX8_9POAL	Q6tsx8 bouteloua h	651	39	40.2	318	2	Q84NR4_ORYSA	Q84nr4 oryza sativ
579	39	40.2	268	2	Q6TSX9_9POAL	Q6tsx9 bouteloua h	652	39	40.2	319	2	Q8X9Z7_ECOS7	Q8x9z7 escherichia
580	39	40.2	268	2	Q6TSY0_9POAL	Q6tsy0 bouteloua h	653	39	40.2	320	2	Q8X9Z7_ECOS7	Q8x9z7 escherichia
581	39	40.2	268	2	Q6TSY1_9POAL	Q6tsy1 bouteloua h	654	39	40.2	324	2	Q75QNO_9VIRU	Q75qno bacterioph
582	39	40.2	268	2	Q6TSY2_9POAL	Q6tsy2 bouteloua h	655	39	40.2	325	2	Q68334_VIBCH	Q68334 vibrio chol
583	39	40.2	268	2	Q6TSY3_9POAL	Q6tsy3 bouteloua h	656	39	40.2	325	2	Q8DAY4_VIBUO	Q8day4 vibrio vuln
584	39	40.2	268	2	Q6TSY4_9POAL	Q6tsy4 bouteloua h	657	39	40.2	327	2	Q9KTS1_VIBCH	Q9kts1 vibrio chol
585	39	40.2	268	2	Q6TSY5_9POAL	Q6tsy5 bouteloua h	658	39	40.2	327	2	Q9D9S9_MOUSE	Q9d9s9 mus musculu
586	39	40.2	268	2	Q6TSY6_9POAL	Q6tsy6 bouteloua h	659	39	40.2	329	2	Q9DZC1_9HIV1	Q9dzc1 human immun
587	39	40.2	268	2	Q6TSY7_9POAL	Q6tsy7 bouteloua h	660	39	40.2	329	2	Q9DZC2_9HIV1	Q9dzc2 human immun
588	39	40.2	268	2	Q6TSY8_9POAL	Q6tsy8 bouteloua h	661	39	40.2	330	1	G3PC_LEIME	Q01558 leishmania
589	39	40.2	268	2	Q6TSY9_9POAL	Q6tsy9 bouteloua h	662	39	40.2	331	2	P93795_TRIDA	P93795 tripaacum d
590	39	40.2	268	2	Q6TSY0_9POAL	Q6tsy0 bouteloua h	663	39	40.2	336	1	ASBI_MOUSE	Q9wv74 mus musculu
591	39	40.2	268	2	Q6TAD1_9POAL	Q6tad1 bouteloua t	664	39	40.2	336	1	TS2_MAIZE	P50160 zea mays (m
592	39	40.2	268	2	Q6TAD8_9POAL	Q6tad8 bouteloua d	665	39	40.2	338	2	Q4PI23_USTMA	Q4pi23 ustilago ma
593	39	40.2	268	2	Q6TAD9_9POAL	Q6tad9 bouteloua d	666	39	40.2	341	2	Q58Y29_LACRH	Q58y29 lactobacill
594	39	40.2	268	2	Q6TAE1_9POAL	Q6tae1 bouteloua d	667	39	40.2	341	2	Q58216_LACRH	Q58216 lactobacill
595	39	40.2	268	2	Q6TAE2_9POAL	Q6tae2 bouteloua d	668	39	40.2	341	2	Q58233_LACRH	Q58233 lactobacill
596	39	40.2	268	2	Q6TAE3_9POAL	Q6tae3 bouteloua d	669	39	40.2	342	2	O88083_ENTFA	O88083 enterococcu
597	39	40.2	268	2	Q6TAE4_9POAL	Q6tae4 bouteloua d	670	39	40.2	342	2	O5SIS7_THET8	Q5sis7 thermus the
598	39	40.2	268	2	Q6TAE5_9POAL	Q6tae5 bouteloua d	671	39	40.2	342	2	Q72J58_THET8	Q72j58 thermus the
599	39	40.2	268	2	Q6TAE7_9POAL	Q6tae7 bouteloua d	672	39	40.2	342	2	Q832N1_ENTFA	Q832n1 enterococcu
600	39	40.2	268	2	Q6TAE8_9POAL	Q6tae8 bouteloua d	673	39	40.2	342	2	Q80TI9_MOUSE	Q80ti9 mus musculu
601	39	40.2	268	2	Q6TAE9_9POAL	Q6tae9 bouteloua d	674	39	40.2	349	2	Q9XD92_ECOLI	Q9xd92 escherichia
602	39	40.2	268	2	Q6TAF0_9POAL	Q6taf0 bouteloua d	675	39	40.2	349	2	Q9CJ08_LACLA	Q9cj08 lactococcu
603	39	40.2	268	2	Q6TAF3_9POAL	Q6taf3 bouteloua d	676	39	40.2	349	2	O5R0S2_IDILO	Q5r0s2 idiomarina
604	39	40.2	268	2	Q6TAF4_9POAL	Q6taf4 bouteloua d	677	39	40.2	350	2	Q8GGW5_LACLC	Q8ggw5 lactococcu
605	39	40.2	268	2	Q6TAF5_9POAL	Q6taf5 bouteloua d	678	39	40.2	352	2	O842N5_VIBCH	O842n5 vibrio chol
606	39	40.2	273	2	Q6TAF2_SYMTH	Q6tbf2 symbiobacte	679	39	40.2	356	2	Q5FT89_GLUOK	Q5ft89 gluconobact
607	39	40.2	274	2	Q5WA99_ORYSA	Q5wa99 oryza sativ	680	39	40.2	359	2	O5PZ36_AJECA	O5pzt36 ajellomyces
608	39	40.2	276	2	Q6R4X8_ZEAM	Q6r4x8 zea mays (s	681	39	40.2	363	2	Q5PZ36_AJECA	Q5pzt36 ajellomyces
609	39	40.2	279	2	Q4HYE6_GIBZE	Q4hye6 gibberella	682	39	40.2	364	1	MRW_BORBR	Q7wfr5 bordetella
610	39	40.2	287	2	Q5N4Z4_STNP6	Q5n4z4 synecococc	683	39	40.2	364	1	MRW_BORPA	Q7w4t7 bordetella
611	39	40.2	289	2	Q6R4W5_ZEAMP	Q6r4w5 zea mays (s	684	39	40.2	364	1	MRW_BORPE	Q7vup6 bordetella
612	39	40.2	291	2	Q8PHT6_XANAC	Q8pht6 xanthomonas	685	39	40.2	365	2	Q6C009_YARLI	Q6c009 yarrowia li
613	39	40.2	294	2	Q3VX4_MAIZE	Q3vx4 zea mays (m	686	39	40.2	366	2	Q6CDN8_YARLI	Q6cdn8 yarrowia li
614	39	40.2	294	2	Q93VX5_MAIZE	Q93vx5 zea mays (m	687	39	40.2	368	2	Q6C1S2_YARLI	Q6c1s2 yarrowia li
615	39	40.2	294	2	Q93W09_MAIZE	Q93w09 zea mays (m	688	39	40.2	368	2	Q6C5R4_YARLI	Q6c5r4 yarrowia li

689	39	40.2	368	2	Q6C5X6_YARLI	Q6C5X6 yarrowia li
690	39	40.2	368	2	Q6CBY8_YARLI	Q6CBY8 yarrowia li
691	39	40.2	368	2	Q6CCN0_YARLI	Q6CCN0 yarrowia li
692	39	40.2	368	2	Q6CH50_YARLI	Q6CH50 yarrowia li
693	39	40.2	369	2	Q877S5_BRAJA	Q877S5 bradyrhizob
694	39	40.2	369	2	Q89YD1_BRAJA	Q89YD1 bradyrhizob
695	39	40.2	371	2	Q6CDZ5_YARLI	Q6CDZ5 yarrowia li
696	39	40.2	372	2	Q5WPP2_BACSK	Q5WPP2 bacillus cl
697	39	40.2	373	2	Q8J5N9_SHIFL	Q8J5N9 shigella fl
698	39	40.2	381	2	Q7V9U3_PROMA	Q7V9U3 prochloroco
699	39	40.2	395	2	Q83X16_VIBAN	Q83X16 vibrio angu
700	39	40.2	396	2	Q51EK9_ENTHI	Q51EK9 entamoeba h
701	39	40.2	403	2	Q7Y0B1_ORYSA	Q7Y0B1 oryza sativ
702	39	40.2	404	2	Q6N8A6_RHOPA	Q6N8A6 rhodopseudo
703	39	40.2	405	2	Q5PZ38_AJECA	Q5PZ38 ajellomyces
704	39	40.2	410	2	Q9LMR4_ARATH	Q9LMR4 arabidopsis
705	39	40.2	414	2	Q4L061_9HIV1	Q4L061 human immun
706	39	40.2	426	2	Q6SGV9_9BACT	Q6SGV9 uncultured
707	39	40.2	427	2	Q6KZ55_PICTO	Q6KZ55 picophilus
708	39	40.2	429	1	SYN_THEAC	Q9H8E7 thermoplas
709	39	40.2	429	1	SYN_THEVO	Q979Y4 thermoplas
710	39	40.2	429	2	Q7S2Z7_NEUCR	Q7S2Z7 neurospora
711	39	40.2	438	2	Q5M9U2_TOBAC	Q5M9U2 nicotiana t
712	39	40.2	441	2	Q87HJ3_VIBPA	Q87HJ3 vibrio para
713	39	40.2	442	2	Q6W4T1_VIBAN	Q6W4T1 vibrio angu
714	39	40.2	445	2	Q70WP6_AERSA	Q70WP6 aeromonas s
715	39	40.2	448	2	Q6CM15_KLULA	Q6CM15 kluyveromyc
716	39	40.2	462	2	Q5BP96_AEDAE	Q5BP96 aedes aegypt
717	39	40.2	465	2	Q6ZFM2_ORYSA	Q6ZFM2 oryza sativ
718	39	40.2	471	2	Q7W4T3_BORPA	Q7W4T3 bordetella
719	39	40.2	473	2	Q7PKB4_ANOGA	Q7PKB4 anopheles g
720	39	40.2	474	2	Q7VU97_BORBR	Q7VU97 bordetella
721	39	40.2	474	2	Q7WGA9_BORBR	Q7WGA9 bordetella
722	39	40.2	475	2	Q6KAQ6_USTMA	Q6KAQ6 mus musculu
723	39	40.2	481	2	Q4P4B5_USTMA	Q4P4B5 ustilago ma
724	39	40.2	486	2	Q6EYV2_YERPS	Q6EYV2 versinia ps
725	39	40.2	489	1	OCLN_POTTR	Q28793 potorous tr
726	39	40.2	491	2	Q5V4L8_HALMA	Q5V4L8 haloarcuula
727	39	40.2	492	2	Q8USW7_9HIV1	Q8USW7 human immun
728	39	40.2	504	2	Q8GW59_ARATH	Q8GW59 arabidopsis
729	39	40.2	508	2	Q6J5N2_9MYRI	Q6J5N2 spheroother
730	39	40.2	516	2	Q5CW26_CRYPV	Q5CW26 cryptospori
731	39	40.2	516	2	Q5CG37_CRYHO	Q5CG37 cryptospori
732	39	40.2	519	2	Q5XIS3_RAT	Q5XIS3 rattus norv
733	39	40.2	531	2	Q7U8T8_SYNPX	Q7U8T8 synechococc
734	39	40.2	533	2	Q9KGJ0_BACHD	Q9KGJ0 bacillus ha
735	39	40.2	537	2	Q91F86_IRV6	Q91F86 chilo iride
736	39	40.2	539	2	Q5UYP5_HALMA	Q5UYP5 haloarcuula
737	39	40.2	550	2	Q5B1T8_EMENI	Q5B1T8 aspergillus
738	39	40.2	551	2	Q5Z8N9_ORYSA	Q5Z8N9 oryza sativ
739	39	40.2	551	2	Q8SB30_ORYSA	Q8SB30 oryza sativ
740	39	40.2	554	2	Q870V7_NEUCR	Q870V7 neurospora
741	39	40.2	562	2	Q9W409_DROME	Q9W409 drosophila
742	39	40.2	567	2	Q27090_TRIVA	Q27090 trichomonas
743	39	40.2	567	2	Q27093_TRIVA	Q27093 trichomonas
744	39	40.2	571	1	DNA43_YEAST	P32354 saccharomyc
745	39	40.2	577	2	Q63X25_BURPS	Q63X25 burkholderi
746	39	40.2	580	2	Q82DF0_STRAW	Q82DF0 streptomyce
747	39	40.2	584	2	Q84DD9_VIBPA	Q84DD9 vibrio para
748	39	40.2	584	2	Q87FY4_VIBPA	Q87FY4 vibrio para
749	39	40.2	591	2	Q5XIT5_RAT	Q5XIT5 rattus norv
750	39	40.2	592	2	Q4ICT8_GIBZE	Q4ICT8 gibberella
751	39	40.2	607	2	Q73HG5_WOLFM	Q73HG5 wolbachia p
752	39	40.2	609	2	Q9JMM7_9RICK	Q9JMM7 9orick
753	39	40.2	609	2	Q60CU6_METCA	Q60CU6 methylococc
754	39	40.2	610	2	Q9JMP6_METCA	Q9JMP6 bacterioph
755	39	40.2	619	2	Q4KKX0_HUMAN	Q4KKX0 homo sapien
756	39	40.2	644	1	VE1_HPV58	Q69JZ1 oryza sativ
757	39	40.2	658	2	Q8D3P3_VIBVU	P26543 human papil
758	39	40.2	664	2	Q8D3P3_VIBVU	Q9BNW5 tomocerus s
759	39	40.2	664	2	Q8D3P3_VIBVU	Q8D3P3 vibrio vuln
760	39	40.2	664	2	Q7MF63_VIBVY	Q7MF63 vibrio vuln
761	39	40.2	690	1	EPB42_HUMAN	P16452 homo sapien

Q4VB97	homo sapien	Q4VB97_HUMAN
Q663C4	versinia ps	Q663C4_YERPS
P14400	torpedo cal	ENP1_TORCA
Q6JUC0	abacion mag	Q6JUC0_9MYRI
Q6BN04	debariomyce	Q6BN04_DEBHA
Q5NHQ1	franciselila	Q5NHQ1_FRATT
Q7XQR1	oryza sativ	Q7XQR1_ORYSA
Q9AUI2	arabidopsis	Q9AUI2_ARATH
Q6UUI4	oryza sativ	Q6UUI4_ORYSA
Q7QE40	anopheles g	Q7QE40_ANOGA
Q6PBA6	brachydanio	Q6PBA6_BRARE
Q9WJ2	drosophila	Q9WJ2_DROME
O04095	arabidopsis	O04095_ARATH
Q42497	chlorella v	Q42497_CHLVU
Q4Z52	pseudomonas	Q4Z52_PSES
Q881L7	pseudomonas	Q881L7_PSES
Q7UN41	rhodopirell	Q7UN41_RHOBA
Q9VJ04	drosophila	Q9VJ04_DROME
Q8S5C7	oryza sativ	Q8S5C7_ORYSA
O18115	caenorhabdi	PSD1_CAREL
Q60K17	caenorhabdi	Q60K17_CABBR
Q52G37	magnaporthe	Q52G37_MAGGR
Q8C8Q8	drosophila	Q8C8Q8_DROME
Q9W495	drosophila	Q9W495_DROME
Q8A5V8	bacteroides	Q8A5V8_BACTN
Q8CGA7	mus musculu	Q8CGA7_MOUSE
Q4Q7I3	leishmania	Q4Q7I3_LEIMA
Q8JAI0	chimpanzee	Q8JAI0_SIVCZ
Q8VH17	mus musculu	Q8VH17_MOUSE
Q8V1K5	mus musculu	Q8V1K5_MOUSE
Q7PVU9	anopheles g	Q7PVU9_ANOGA
Q88099	chimpanzee	Q88099_SIVCZ
P16462	actinobacil	LKTA_ACTAC
Q43892	actinobacil	Q43892_ACTAC
Q7UP09	rhodopirell	Q7UP09_RHOBA
P89904	chimpanzee	P89904_SIVCZ
O8JAV1	vienna/naedi	O8JAV1_9RETR
Q561B0	cryptococcu	Q561B0_CRYNE
Q4IPF9	gibberella	Q4IPF9_GIBZE
Q5KQ11	cryptococcu	Q5KQ11_CRYNE
Q4P4F6	ustilago ma	Q4P4F6_USTMA
Q560T0	cryptococcu	Q560T0_CRYNE
Q5KPS9	cryptococcu	Q5KPS9_CRYNE
Q54Y47	dictyosteli	Q54Y47_DICDI
Q9USC7	schizosacch	Q9USC7_SCHPO
Q8WJ61	brachydanio	Q8WJ61_BRARE
Q8WJ61	homo sapien	CEND3_HUMAN
Q7FAP7	oryza sativ	Q7FAP7_ORYSA
Q8IAZ4	homo sapien	Q8IAZ4_HUMAN
Q8SA91	zea mayz	Q8SA91_MAZE
Q6B144	debariomyce	Q6B144_DEBHA
Q6J466	toxoplasma	Q6J466_TOXGO
Q8T5B3	actinobact	Q6FE11_ACTAD
Q8T5B3	metanopyru	KCY_METKA
Q9FC61	streptomyce	Q9FC61_STRGO
Q6G741	staphylococ	Q6G741_STAAS
Q5HDS6	staphylococ	Q5HDS6_STAAC
Q7AA36	staphylococ	Q7AA36_STAAN
Q8NV94	staphylococ	Q8NV94_STAAM
Q99Z21	staphylococ	Q99Z21_STAAM
Q63U03	burkholderi	Q63U03_BURPS
Q6AEX6	leifsonia x	Q6AEX6_LEIXX
Q6Z1J7	oryza sativ	Q6Z1J7_ORYSA
P52322	microcystis	RPD0_MICAE
Q9P5Z3	neurospora	Q9P5Z3_NEUCR
Q7Q159	anopheles g	Q7Q159_ANOGA
Q5TP03	anopheles g	Q5TP03_ANOGA
Q96MU6	homo sapien	Q96MU6_HUMAN
Q9G8W2	drosophila	Q9G8W2_DROME
Q61K83	caenorhabdi	Q61K83_CABBR
Q7RUZ6	neurospora	Q7RUZ6_NEUCR
Q6R573	trichoderma	Q6R573_9HYPO
Q9G8W1	drosophila	Q9G8W1_DROVI

835	38.5	39.7	516	2	Q6M3H3_CORGL	Q6m3h3 corynebacte	908	38	39.2	228	2	Q9VRA4_DROME	Q9vra4 drosophila
836	38.5	39.7	519	2	Q6C3N9_YARLI	Q6c3n9 yarrowia li	909	38	39.2	228	2	Q8SZE4_DROME	Q8sze4 drosophila
837	38.5	39.7	520	2	Q8NND9_CORGL	Q8nnd9 corynebacte	910	38	39.2	229	2	Q8KM06_ENTFA	Q8km06 enterococcu
838	38.5	39.7	529	2	Q57IL8_SALCH	Q57il8 salmonella	911	38	39.2	230	2	Q88T17_LACPL	Q88t17 lactobacill
839	38.5	39.7	529	2	Q5FUQ0_SALPA	Q5fuq0 salmonella	912	38	39.2	230	2	Q72KL1_9HIV1	Q72kl1 human immun
840	38.5	39.7	529	2	Q82Z78_SALTI	Q82z78 salmonella	913	38	39.2	241	2	Q21302_CABEL	Q21302 caenorhabdi
841	38.5	39.7	529	2	Q8ZLC7_SALTY	Q8zlc7 salmonella	914	38	39.2	241	2	Q66269_9SPHN	Q66269 erythromicr
842	38.5	39.7	559	2	Q9GSM3_DROVI	Q9gsm3 drosophila	915	38	39.2	245	2	Q66278_9SPHN	Q66278 agrobacteri
843	38.5	39.7	587	2	Q5CLN1_CRYPV	Q5cln1 cryptospori	916	38	39.2	245	2	Q66278_9SPHN	Q66278 agrobacteri
844	38.5	39.7	587	2	Q5CLN1_CRYHO	Q5cln1 cryptospori	917	38	39.2	246	2	Q66276_9SPHN	Q66276 porphyrobac
845	38.5	39.7	639	2	Q7NRI4_CHRVO	Q7nri4 chromobacte	918	38	39.2	246	2	Q66276_9SPHN	Q66276 porphyrobac
846	38.5	39.7	659	2	Q4J480_AZOV1	Q4j480 azotobacter	919	38	39.2	247	2	Q24659_ORYSA	Q24659 oryza sativ
847	38.5	39.7	665	2	Q870B6_9FUNG	Q870b6 piromyces s	920	38	39.2	247	2	Q8S1X5_ORYSA	Q8s1x5 oryza sativ
848	38.5	39.7	857	2	Q24736_DROVI	Q24736 drosophila	921	38	39.2	250	2	Q88VQ0_LACPL	Q88vq0 lactobacill
849	38.5	39.7	857	2	Q6S015_DROME	Q6s015 drosophila	922	38	39.2	250	2	Q24648_BETVU	Q24648 beta vulgar
850	38.5	39.7	857	2	Q6S016_DROME	Q6s016 drosophila	923	38	39.2	250	2	Q88VQ0_LACPL	Q88vq0 lactobacill
851	38.5	39.7	857	2	Q9SVBS_DROVI	Q9svbs drosophila	924	38	39.2	250	2	Q24648_BETVU	Q24648 beta vulgar
852	38.5	39.7	1073	1	RAGL_ONCMY	Q91187 oncorhynchu	925	38	39.2	250	2	Q88VQ0_LACPL	Q88vq0 lactobacill
853	38.5	39.7	2476	1	ZAN_PIG	Q28983 sus scrofa	926	38	39.2	254	1	Q11777_9PARA	Q11777 measles vir
854	38	39.2	31	2	Q4YGC2_PLABE	Q4ygc2 plasmodium	927	38	39.2	254	1	Q11777_9PARA	Q11777 measles vir
855	38	39.2	68	2	Q7R2G9_GIALA	Q7r2g9 giardia lam	928	38	39.2	256	2	Q11777_9PARA	Q11777 measles vir
856	38	39.2	68	2	Q6XK08_SPICI	Q6xk08 spiroplasma	929	38	39.2	256	2	Q07942_BACILLU	Q07942 bacillus su
857	38	39.2	84	2	Q6C8D7_YARLI	Q6c8d7 yarrowia li	930	38	39.2	257	2	Q03419_HUMAN	Q03419 human respi
858	38	39.2	99	2	Q4HKA2_CAMLA	Q4hka2 campylobact	931	38	39.2	257	2	Q77YB3_HUMAN	Q77yb3 human respi
859	38	39.2	101	2	Q5ZC53_ORYSA	Q5zc53 oryza sativ	932	38	39.2	260	2	Q4KRW7_HRSV	Q4krw7 human respi
860	38	39.2	102	2	Q5L786_CHLAB	Q5l786 chlamydophi	933	38	39.2	261	2	Q92948_CHLPL	Q92948 chlamydia p
861	38	39.2	104	2	Q92U58_RHIME	Q92u58 rhizobium m	934	38	39.2	262	2	Q49657_MYCLE	Q49657 mycobacteri
862	38	39.2	119	2	Q7SPR6_NEUCR	Q7spr6 neurospora	935	38	39.2	262	2	Q5IR78_9HIV1	Q5ir78 human immun
863	38	39.2	125	2	Q5QNA1_ORYSA	Q5qna1 oryza sativ	936	38	39.2	279	2	Q9YQ24_SIVCZ	Q9yq24 chimpanzee
864	38	39.2	134	2	Q9RW17_DETRA	Q9rw17 deinococcus	937	38	39.2	281	2	Q6IGU0_DROME	Q6igu0 drosophila
865	38	39.2	135	2	Q4TPZ6_9SPHN	Q4tpz6 erythrobact	938	38	39.2	282	2	Q6IGU0_DROME	Q6igu0 drosophila
866	38	39.2	136	2	Q9DAL0_MOUSE	Q9dal0 mus musculu	939	38	39.2	285	2	Q62PT4_MOUSE	Q62pt4 mus musculu
867	38	39.2	140	2	Q8EUP0_MYCPE	Q8eup0 mycoplasma	940	38	39.2	285	2	Q9HLI1_THRAC	Q9hli1 thermoplasma
868	38	39.2	147	2	Q82Y81_NITEU	Q82y81 nitrosomona	941	38	39.2	285	2	Q5YFP7_9VIRU	Q5yfp7 singapore g
869	38	39.2	151	2	Q4QJQ5_LEIMA	Q4qj05 leishmania	942	38	39.2	285	2	Q5GAL3_9VIRU	Q5gal3 grouper iri
870	38	39.2	153	2	Q6AMX6_DESPS	Q6amx6 desulfofocale	943	38	39.2	286	2	Q7NWB3_CHRVO	Q7nwb3 chromobacte
871	38	39.2	154	2	Q7NZJ2_CHRVO	Q7nzj2 chromobacte	944	38	39.2	286	2	Q6LPY7_PHOPR	Q6lp77 photobacter
872	38	39.2	159	2	Q9AXE8_CITUN	Q9axe8 citrus unsh	945	38	39.2	287	2	Q7VDZ5_PROMA	Q7vdz5 prochloroco
873	38	39.2	160	2	Q73UZ7_MYCPA	Q73uz7 mycobacteri	946	38	39.2	287	2	Q8BV58_MOUSE	Q8bv58 mus musculu
874	38	39.2	163	2	Q83JD4_SHIFL	Q83jd4 shigella fl	947	38	39.2	287	2	Q4NHT3_9MICC	Q4nht3 arthrobacte
875	38	39.2	169	2	Q9ZSW7_HAMVI	Q9zsw7 hamamelis v	948	38	39.2	287	2	Q8FAB5_ECOL6	Q8fab5 escherichia
876	38	39.2	171	2	Q9SBC1_HAMVI	Q9sbc1 hamamelis v	949	38	39.2	301	2	Q8RQF9_9HIV1	Q8rqf9 human immun
877	38	39.2	172	2	Q7RHA0_PLAYO	Q7rha0 plasmodium	950	38	39.2	301	2	Q8EJ00_SHEWANA	Q8ej00 shewanella
878	38	39.2	173	2	Q9SBC0_HAMVI	Q9sbc0 hamamelis v	951	38	39.2	306	2	Q83NN4_TROW8	Q83nn4 tropheryma
879	38	39.2	173	2	Q9ZSW6_HAMVI	Q9zsw6 hamamelis v	952	38	39.2	309	2	Q73H24_WOLPM	Q73h24 wolbachia p
880	38	39.2	175	2	Q5U648_HUMAN	Q5u648 homo sapien	953	38	39.2	311	1	LYSR_ECOLI	LYSR escherichia
881	38	39.2	179	2	Q5KY62_GEOKA	Q5ky62 geobacillus	954	38	39.2	311	2	Q7AB47_ECO57	Q7ab47 escherichia
882	38	39.2	180	2	Q5V286_HALMA	Q5v286 haloarcula	955	38	39.2	311	2	Q57KA3_SALCH	Q57ka3 salmonella
883	38	39.2	181	2	Q9F3Z6_NEIME	Q9f3z6 neisseria m	956	38	39.2	311	2	Q8XGJ2_ECO57	Q8xgj2 escherichia
884	38	39.2	182	2	Q92VU0_RHIME	Q92vu0 rhizobium m	957	38	39.2	311	2	Q8FEAL_ECOL6	Q8feal escherichia
885	38	39.2	183	2	Q4QOC4_LEIMA	Q4qoc4 leishmania	958	38	39.2	311	2	Q7N7G0_PHOLL	Q7n7g0 photorhabdu
886	38	39.2	183	2	Q8D7J6_VIBVU	Q8d7j6 vibrio vuln	959	38	39.2	311	2	Q7CPV5_SALTY	Q7cpv5 salmonella
887	38	39.2	184	2	Q4NNK9_9DELT	Q4nnk9 anaeromyxob	960	38	39.2	311	2	Q5PEP7_SALPA	Q5pep7 salmonella
888	38	39.2	185	2	Q975X8_SULTO	Q975x8 sulfolobus	961	38	39.2	312	2	Q8XGD5_SALTI	Q8xgd5 salmonella
889	38	39.2	190	2	Q7Q0R7_ANOGA	Q7q0r7 anopheles g	962	38	39.2	312	2	Q8XGD5_SALTI	Q8xgd5 shigella fl
890	38	39.2	190	2	Q6DT74_ARALP	Q6dt74 arabisopsis	963	38	39.2	312	2	Q5PT51_MOUSE	Q5pt51 mus musculu
891	38	39.2	191	2	Q9RDX0_STRCO	Q9rjx0 streptomyce	964	38	39.2	312	2	Q667E7_YERPS	Q667e7 yersinia ps
892	38	39.2	194	1	Y004_MYCSM	Q50434 mycobacteri	965	38	39.2	312	2	Q83MV4_TROWT	Q83mv4 tropheryma
893	38	39.2	196	2	Q8SAP5_9BRYO	Q8sap5 mitthyridiu	966	38	39.2	316	2	Q83MV4_TROWT	Q83mv4 tropheryma
894	38	39.2	198	1	RECR_GLUOX	Q5ftb6 gluconobact	967	38	39.2	320	2	Q982F1_RHILO	Q982f1 rhizobium l
895	38	39.2	199	2	Q9X6K1_KLEOX	Q9x6k1 klebsiella	968	38	39.2	320	2	Q39288_9ALPH	Q39288 equid herpe
896	38	39.2	203	1	IFNB_CHICK	Q9x6k1 klebsiella	969	38	39.2	322	2	Q9PTB4_BRARE	Q9ptb4 brachydanio
897	38	39.2	203	2	Q5MFU9_CHICK	Q9x6k1 klebsiella	970	38	39.2	327	1	Y745_HELPU	Y745 helicobacte
898	38	39.2	206	2	Q622A8_CAEBR	Q622a8 caenorhabdi	971	38	39.2	327	2	Q5H9S6_HUMAN	Q5h9s6 homo sapien
899	38	39.2	208	2	Q51DP1_ENTHI	Q51dp1 entamoeba h	972	38	39.2	328	2	Q8CZX8_YERPE	Q8czx8 yersinia pe
900	38	39.2	209	2	Q5N2P0_SYN6P	Q5n2p0 synecococc	973	38	39.2	331	2	Q9YAS2_AERPE	Q9yas2 aeropyrum p
901	38	39.2	210	2	Q6IGL5_DROME	Q6igl5 drosophila	974	38	39.2	331	2	Q6NJB6_CORDI	Q6njb6 corynebacte
902	38	39.2	214	2	Q8DD88_VIBVU	Q8dde8 vibrio vuln	975	38	39.2	333	1	Q5ZJH8_CHICK	Q5zjh8 gallus gall
903	38	39.2	214	2	Q7MGR1_VIBVY	Q7mgi1 vibrio vuln	976	38	39.2	335	1	Y481_STAAS	Y481 staphylococ
904	38	39.2	222	2	Q9U796_MUSDO	Q9u796 musca domes	977	38	39.2	335	1	Y524_STAAM	Y524 staphylococ
905	38	39.2	223	2	Q9XJ28_ORYSA	Q9xj28 oryza sativ	978	38	39.2	335	2	Q4JDU2_9PROT	Q4jdu2 nitrosomona
906	38	39.2	225	2	Q7M8L3_WOLSU	Q7m8l3 wolinnella s	979	38	39.2	335	2	Q5HIF0_STAAC	Q5hif0 staphylococ
907	38	39.2	227	2	Q8NBP8_HUMAN	Q8nbp8 homo sapien	980	38	39.2	336	1	SMT1_ARATH	SMT1 arabidopsis

Q9vra4 drosophila	Q8sze4 drosophila	Q8km06 enterococcu	Q72kl1 human immun	Q21302 caenorhabdi	Q66269 erythromicr	Q66278 agrobacteri	Q66278_9SPHN	Q66276_9SPHN	Q24659_ORYSA	Q8S1X5_ORYSA	Q88VQ0_LACPL	Q24648_BETVU	Q88VQ0_LACPL	Q24648_BETVU	Q11777_9PARA	Q11777_9PARA	Q07942_BACILLU	P03419_HUMAN	Q77YB3_HUMAN	Q4KRW7_HRSV	Q92948_CHLPL	Q49657_MYCLE	Q5IR78_9HIV1	Q9YQ24_SIVCZ	Q6FY71_ARATH	Q6IGU0_DROME	Q62PT4_MOUSE	Q9HLI1_THRAC	Q5YFP7_9VIRU	Q5GAL3_9VIRU	Q7NWB3_CHRVO	Q6LPY7_PHOPR	Q7VDZ5_PROMA	Q8BV58_MOUSE	Q4NHT3_9MICC	Q8FAB5_ECOL6	Q8RQF9_9HIV1	Q8EJ00_SHEWANA	Q83NN4_TROW8	Q73H24_WOLPM	LYSR_ECOLI	Q7AB47_ECO57	Q57KA3_SALCH	Q8XGJ2_ECO57	Q8FEAL_ECOL6	Q7N7G0_PHOLL	Q7CPV5_SALTY	Q5PEP7_SALPA	Q8XGD5_SALTI	Q83MV4_TROWT	Q982F1_RHILO	Q39288_9ALPH	Q9PTB4_BRARE	Y745_HELPU	Q5H9S6_HUMAN	Q8CZX8_YERPE	Q9YAS2_AERPE	Q6NJB6_CORDI	Q5ZJH8_CHICK	Y481_STAAS	Y524_STAAM	Q4JDU2_9PROT	Q5HIF0_STAAC	SMT1_ARATH
-------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	----------------	--------------	--------------	-------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	----------------	--------------	--------------	------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	------------	--------------	--------------	--------------	--------------	--------------	------------	------------	--------------	--------------	------------

981 38 39.2 336 2 Q5V542 HALMA
982 38 39.2 336 2 Q8A137 helicobacte
983 38 39.2 337 1 G3F PHACH
984 38 39.2 340 2 Q5PT50 RAT
985 38 39.2 340 2 Q9CWD6 MOUSE
986 38 39.2 340 2 Q5PT53 MOUSE
987 38 39.2 342 1 HEM2 YEAST
988 38 39.2 343 2 Q7T0V6 XENLA
989 38 39.2 346 2 Q24328 RICCO
990 38 39.2 350 2 Q677X3 VIRU
991 38 39.2 351 2 Q5AF67 CANAL
992 38 39.2 353 2 Q412H3 GIBZE
993 38 39.2 354 2 Q8ZVT5 PYRAE
994 38 39.2 354 2 Q8Y094 RALSO
995 38 39.2 356 2 Q424G6 FLABE
996 38 39.2 356 2 Q84138 HELPY
997 38 39.2 356 2 Q9A2D6 CAUCR
998 38 39.2 356 2 Q6JG27 9HIV1
999 38 39.2 357 2 Q6GAB2 STAAS
1000 38 39.2 357 2 Q6GHY3 STAAR

ALIGNMENTS

RESULT 1
ID Q8ZC84_YERPE PRELIMINARY; PRT; 122 AA.
AC Q8ZC84_Q74WQ2_Q7CK17;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Hypothetical protein YPO3137 (Hypothetical protein Y1047).
GN OrderedLocusNames=YPO794, YPO3137, Y1047;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parthill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatina M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Farraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Zhou D., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Fang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
avirulent to humans."

RL DNA Res. 11:179-197(2004).
DR EMBL: AJ414155; CAC92372.1; -; Genomic_DNA.
DR EMBL: AE013708; NAM84628.1; -; Genomic_DNA.
DR EMBL: AE017129; AAS61059.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 14215 MW; 7C49F0A1B8EBC157 CRC64;
Query Match 50.5%; Score 49; DB 2; Length 122;
Best Local Similarity 53.3%; Pred. No. 6.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 IKGPTLRQWLKSRH 16
DB 92 INDPRLRRWQTKKH 106
RESULT 2
ID Q66DR4_YERPS PRELIMINARY; PRT; 122 AA.
AC Q66DR4;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=YPTB0579;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype J;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL: BX936398; CAH20219.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 122 AA; 14215 MW; 7C49F0A1B8EBC157 CRC64;
Query Match 50.5%; Score 49; DB 2; Length 122;
Best Local Similarity 53.3%; Pred. No. 6.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 IKGPTLRQWLKSRH 16
DB 92 INDPRLRRWQTKKH 106
RESULT 3
ID Q5AQW2_EMENI PRELIMINARY; PRT; 718 AA.
AC Q5AQW2;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AN9318.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,

```

RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepl Y., Collymore A., Cook K., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kelle C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Maucelli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramaamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans";
CC Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC ENBL; AACD01000172; EAA66385.1; -; Genomic_DNA.
KW Hypochemical protein.
SQ SEQUENCE 718 AA; 79121 MW; 9E865DCECC4DE39 CRC64;

Query Match 50.0%; Score 48.5; DB 2; Length 718;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 5 PTL-RQWLKSRHENTS 18
||| :||| :||| :
Db 478 PTLRKWLKPRANTA 492

RESULT 4
Q4UK00_RICFE PRELIMINARY; PRT; 148 AA.
AC Q4UK00;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Uncharacterized conserved protein.
GN OrderedLocusNames=RF1288;
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia; spotted fever group.
OX NCBI_TaxID=42862;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=UPRWKCal2.
RX PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA Ogata H., Renesto P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA Parinello H., Claverie J.-M., Raoult D.;
RT "The genome sequence of Rickettsia felis identifies the first putative
RT conjugative plasmid in an obligate intracellular parasite.";
RL PLoS Biol. 3:E248-E248(2005).
DR EMBL; CP000053; AAY62139.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 148 AA; 17278 MW; 98592DA61DDEE520 CRC64;

Query Match 49.5%; Score 48; DB 2; Length 148;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWL 11
||||| :|||
Db 70 TIKGFSSKQWL 80

```

```

RESULT 5
P95613_RHIGA PRELIMINARY; PRT; 326 AA.
ID P95613_RHIGA PRELIMINARY; PRT; 326 AA.
AC P95613;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NodD2 protein.
OS Name=nodD2;
GN Rhizobium galegae.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HAMBI;
RA Suominen L., Roos C., Paulin L., Kaijalainen S., Lindstroem K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH LYSR-type DNA-binding domain.
CC EMBL; Y08963; CAA70157.1; -; Genomic_DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000847; HTH_LYSR.
DR InterPro; IPR005119; LysR_subst_bd.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF01266; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS50931; HTH_LYSR; 1.
DR Activator; DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 326 AA; 36373 MW; BFE9C32F6719E28B CRC64;

Query Match 49.5%; Score 48; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSR 14
||||| :||| :
Db 204 KGFSLEQWLSSQ 215

RESULT 6
Q4P0J6_USTMA PRELIMINARY; PRT; 992 AA.
ID Q4P0J6_USTMA PRELIMINARY; PRT; 992 AA.
AC Q4P0J6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM06367.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilago.
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitsheteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Duffey N., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamvyselis M., Karlsson E.,

```

```

RA Kelle C., Kieu A., Kianer P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Landlad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Macbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrum J., Meneus L.,
RA Mesirov J., Mihalav A., Mihova T., Mikelsen T., Mienga V., Moru K.,
RA Mozes J., Mullrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okawo O., O'leary S., Omotosho B.,
RA O'Neill K., Oman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tensing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Teamlu T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Zander E.;
RT "The genome sequence of Ustilago maydis."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC00100247; EAK87224.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 992 AA; 106443 MW; 2F60D4BA01A2011F CRC64;

Query Match 49.5%; Score 48; DB 2; Length 992;
Best Local Similarity 57.1%; Pred. No. 94;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRWLKSR 14
Db 561 TVGGPALRMWRKAR 574

RESULT 7
Y745_HELPJ STANDARD; PRT; 327 AA.
AC Q9ZL38;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical RNA pseudouridine synthase JHP0682 (EC 5.4.99.-) (RNA-
DE uridine isomerase) (RNA pseudouridylylate synthase).
GN OrderedLocuNames-JHP0682;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180 (1999).
CC -!- CATALYTIC ACTIVITY: RNA uridine = RNA pseudouridine.
CC -!- SIMILARITY: Belongs to the pseudouridine synthase rluA family.
CC -!- SIMILARITY: Contains 1 S4 RNA-binding domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AE001500; AAD06270.1; -; Genomic_DNA.
DR F1R; B71900; B71900.
DR InterPro; IPR006225; Pseud_synth_RluD.
DR InterPro; IPR006145; Pseudou_synth.
DR InterPro; IPR006224; Rlu_synth.
DR InterPro; IPR002942; S4_RNA_bd.
DR Pfam; PF00849; Pseudou_synth_2; 1.
DR Pfam; PF01479; S4; 1.
DR ProDom; PD001819; Pseudou_synth; 1.
DR SMART; SM00363; S4; 1.
DR TIGRFAMs; TIGR00005; rluD_subfam; 1.
DR PROSITE; PS01129; PSI_RLU; 1.
DR PROSITE; PS50889; S4; 1.
KW Complete proteome; Hypothetical protein; Isomerase; RNA-binding.
FT DOMAIN 12 79 S4 RNA-binding.
FT ACT_SITE 136 136 By similarity.
SQ SEQUENCE 327 AA; 37722 MW; 7EDC7F6840D818BD CRC64;

Query Match 48.5%; Score 47; DB 1; Length 327;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 TIKGPTLRWLKSR 16
Db 103 SVKEPTLVDLKSNQY 118

RESULT 8
Q750V6_ASHGO PRELIMINARY; PRT; 727 AA.
AC Q750V6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE AGL167Cp.
GN Name=AGL167C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehmann R., Luedi P., Choi S., Wing R.A., Flavien A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome."
RL Science 304:304-307 (2004).
DR EMBL; AF016820; AAS54324.1; -; Genomic_DNA.
DR AGD; AGL167C; -.
KW Complete proteome.
SQ SEQUENCE 727 AA; 82748 MW; 58A66322705F6767 CRC64;

Query Match 48.5%; Score 47; DB 2; Length 727;
Best Local Similarity 56.2%; Pred. No. 97;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KGFTLRWLKSRHTS 18
Db 92 RGEGTSWRKDRHGS 107

RESULT 9
Q55ZAJ_CRYNE PRELIMINARY; PRT; 1794 AA.
ID Q55ZAJ_CRYNE
AC Q55ZAJ;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)

```

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBA5860;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AA01000004; EAL23061.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PSS0003; PH_DOMAIN; 1.
DR PROSITE; PSS0190; SEC7; 1.
KW Coiled coil; Guanine-nucleotide releasing factor;
KW Hypothetical protein.
SQ SEQUENCE 1794 AA; 196453 MW; 98811FD2F194C994 CRC64;
Query Match 48.5%; Score 47; DB 2; Length 1794;
Best Local Similarity 46.2%; Pred. No. 2.6e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSRH 16
DB 589 GPGIRPMMETROH 601
RESULT 10
QSKNL7 CRYNE PRELIMINARY; PRT; 1811 AA.
AC QSKNL7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE AAF guanyl-nucleotide exchange factor, putative.
GN ORFNames=CNBA06050;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bodet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Krzywinski M.I., Kwon-Chung J.K.,
RA Langelier K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA Davis R.W., Fraser C.M., Hyman R.W.;
RT "The genome and transcriptome of Cryptococcus neoformans, a
RT basidiomycete fungal pathogen of humans."
RL Science 0:0-0(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bodet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung J.K., Langelier K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans."
RL Science 307:1321-1324(2005).
CC -!- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB017341; AAW41140.1; -; Genomic_DNA.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR000904; Sec7.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01369; Sec7; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00222; Sec7; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PSS0003; PH_DOMAIN; 1.
DR PROSITE; PSS0190; SEC7; 1.
KW Coiled coil; Complete proteome; Guanine-nucleotide releasing factor.
SQ SEQUENCE 1811 AA; 198207 MW; 33888A8C9C924019 CRC64;
Query Match 48.5%; Score 47; DB 2; Length 1811;
Best Local Similarity 46.2%; Pred. No. 2.7e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 4 GPTLRQWLKSRH 16
DB 606 GPGIRPMMETROH 618
RESULT 11
GPDA2 MYCPA
ID GPDA2 MYCPA STANDARD; PRT; 332 AA.
AC P61744;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Glycerol-3-phosphate dehydrogenase 2 [NAD(P)+] [EC 1.1.1.94] (NAD(P)H-
DE dependent glycerol-3-phosphate dehydrogenase 2).
GN Name-gpsA2; OrderedLocNames=MAP4061c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: sn-glycerol 3-phosphate + NAD(P)(+) =
CC glycerone phosphate + NAD(P)H.
CC -!- PATHWAY: De novo phospholipid biosynthesis; glycerol-3 phosphate
CC formation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the NAD-dependent glycerol-3-phosphate

```
CC      dehydrogenase family.
CC      -----
CC      This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      EMBL; AB017241; AA050611.1; -; Genomic_DNA.
CC      HAMAP; MF_00394; -; 1.
CC      InterPro; IPR006109; NAD_Gly3P_C.
CC      InterPro; IPR006168; NAD_Gly3P_dh.
CC      InterPro; IPR011128; NAD_Gly3P_dh_N.
CC      PANTHER; PTHR111728; NAD_Gly3P_dh; 1.
CC      Pfam; PF07479; NAD_Gly3P_dh_C; 1.
CC      Pfam; PF01210; NAD_Gly3P_dh_N; 1.
CC      PRINTS; PR00077; GPDHGRNASE.
CC      PRODOM; PD001278; NAD_Gly3P_C; 1.
CC      PROSITE; PS00957; NAD_G3PDH; 1.
CC      Complete proteome; NAD; Oxidoreductase; Phospholipid biosynthesis.
CC      SEQUENCE 332 AA; 35220 MW; B149EF540B313D83 CRC64;
SQ
Query Match 47.9%; Score 46.5; DB 1; Length 332;
Best Local Similarity 69.2%; Pred. No. 49;
Matches 9; Conservative 3; Mismatches 0; Indels 1; Gaps 1;
QY      3 KGPTLRQWLKRSRE 15
DB      29 RGPTL-QWVRSRE 40
      :|||||
      -|- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid
      dehydrogenase family.
RESULT 12
Q6YUC3 ORYSA PRELIMINARY; PRT; 126 AA.
AC      Q6YUC3
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE      Hypothetical protein B1116H04.2 (Hypothetical protein
DE      B1111C03.14).
GN      Name=B1116H04.2; Synonym=B1111C03.14;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzaeae; Oryza.
OX      NCBI_TaxID=39947;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Sasaki T., Matsumoto T., Katayose Y.;
RA      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN      NUCLEOTIDE SEQUENCE.
RA      Sasaki T., Matsumoto T., Katayose Y.;
RA      Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AP005871; BADI0677.1; -; Genomic_DNA.
DR      EMBL; AP005405; BADI0301.1; -; Genomic_DNA.
DR      Gramene; Q6YUC3; -.
KW      Hypothetical protein.
SQ      SEQUENCE 126 AA; 12558 MW; F316174EC475BAC6 CRC64;
Query Match 47.4%; Score 46; DB 2; Length 126;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY      1 TIKGPTLRQWLKRSRE 15
DB      66 TAAGPATRWVKTRQ 80
      :|||||
      -|- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid
      dehydrogenase family.
RESULT 13
Q6C114 YARLI PRELIMINARY; PRT; 365 AA.
ID      Q6C114 YARLI
AC      Q6C114
```

```
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE      Similar to tr|O93968 Candida boidinii Formate dehydrogenase.
GN      OrderedLocusNames=YAL10F15983g;
OS      Yarrowia lipolytica (Candida lipolytica).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Dipodascaceae; Yarrowia.
OX      NCBI_TaxID=4952;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=CLIB 122 / E 150;
RX      PubMed=15229592; DOI=10.1038/nature02579;
RA      Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA      Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA      Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA      Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleksten C.,
RA      Boirame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA      Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA      Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA      Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA      Niclaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA      Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA      Swennen D., Tekaita F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA      Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA      Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA      Wincker P., Souciet J.-L.;
RA      "Genome evolution in yeasts.";
RL      Nature 430:35-44(2004).
CC      -|- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid
CC      dehydrogenase family.
CC      EMBL; CR382132; CAG78287.1; -; Genomic DNA.
CC      GO; GO:0016491; F:oxidoreductase activity; IEA.
CC      GO; GO:0016616; F:oxidoreductase activity, acting on the CH-O. .; IEA.
CC      GO; GO:0006564; P:L-serine biosynthesis; IEA.
CC      InterPro; IPR006139; 2-Hacid_DH.
CC      InterPro; IPR006140; 2-Hacid_DH_C.
CC      Pfam; PF00389; 2-Hacid_dh; 1.
CC      Pfam; PF02826; 2-Hacid_dh_C; 1.
CC      PROSITE; PS00665; D_2_HYDROXYACID_DH_1; 1.
CC      PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
CC      PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
CC      Complete proteome; Oxidoreductase.
CC      SEQUENCE 365 AA; 40172 MW; 8AAC7PB8785139E0 CRC64;
Query Match 47.4%; Score 46; DB 2; Length 365;
Best Local Similarity 72.7%; Pred. No. 66;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      7 LRQWLKSRHT 17
DB      30 LRQWLKSRHT 40
      |||:|||||
      -|- SIMILARITY: Belongs to the D-isomer specific 2-hydroxyacid
      dehydrogenase family.
RESULT 14
Q82R87 STRAW PRELIMINARY; PRT; 402 AA.
ID      Q82R87 STRAW PRELIMINARY;
AC      Q82R87;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Putative transposase.
GN      OrderedLocusNames=SAV256;
OS      Streptomyces avermitilis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Streptomycineae; Streptomycetaceae; Streptomyces.
OX      NCBI_TaxID=33903;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX      MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA      Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA      Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
```



```
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites."; Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RN [2]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis."; Nat. Biotechnol. 21:526-531(2003).  
RL Nat. Biotechnol. 21:526-531(2003).  
DR EMBL; BA000030; BAC67965.1; -: Genomic_DNA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0004803; F:transposase activity; IEA.  
DR GO; GO:0006313; P:DNA transposition; IEA.  
DR InterPro; IPR002559; Transposase_11.  
DR Pfam; PF01609; Transposase_11; 1.  
KW Complete proteome.  
SQ SEQUENCE 402 AA; 43379 MW; 71EC0A91143451BE CRC64;  
  
Query Match 47.4%; Score 46; DB 2; Length 402;  
Best Local Similarity 71.4%; Pred. No. 73;  
Matches 10; Conservative 1; Mismatches 1; Indels 2; Gaps 1;  
  
QY 2 IKG--PTLRWLKS 13  
||| ||| ||| |||  
Db 234 IKGNQPTLHQLKA 247  
  
RESULT 15  
Q4KS46_9VIRU  
ID Q4KS46_9VIRU PRELIMINARY; PRT; 815 AA.  
AC Q4KS46;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Tyrosine kinase.  
OS Orange-spotted grouper iridovirus.  
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
OX NCBI_TaxID=322017;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.  
RA Zhou S.Y., Lv L., Chen C., Weng S.P., Chan S.M., He J.G.;  
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY894343; AAX82420.1; -: Genomic_DNA.  
KW Kinase.  
SQ SEQUENCE 815 AA; 93153 MW; 120C2FB59703F8C4 CRC64;  
  
Query Match 47.4%; Score 46; DB 2; Length 815;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 2 IKGPTLRWLKSRHT 17  
:||||| ||| |||  
Db 500 VQGPTLAQWICSTAF 515  
  
RESULT 16  
Q8QUJ6_9VIRU  
ID Q8QUJ6_9VIRU PRELIMINARY; PRT; 941 AA.  
AC Q8QUJ6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE ORF1141.  
OS Infectious spleen and kidney necrosis virus.  
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Megalocytivirus.  
OX NCBI_TaxID=180170;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.
```

```
RX MEDLINE=21974810; PubMed=11878882; DOI=10.1006/viro.2001.1208;  
RA He J.G., Weng S.P., Li Z., Zhou S.Y., Long Q.X., Wang X.Z.,  
RA Chan S.M.;  
RT "Complete genome analysis of the mandarin fish infectious spleen and  
RT kidney necrosis iridovirus"; Virology 291:126-139(2001).  
RL EMBL; AF371960; AAL98838.1; -: Genomic_DNA.  
DR EMBL; AF371960; AAL98838.1; -: Genomic_DNA.  
SQ SEQUENCE 941 AA; 106703 MW; EB663998C7F6CE83 CRC64;  
  
Query Match 47.4%; Score 46; DB 2; Length 941;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 2 IKGPTLRWLKSRHT 17  
:||||| ||| |||  
Db 581 VQGPTLAQWICSTAF 596  
  
RESULT 17  
Q980N7_SULSO  
ID Q980N7_SULSO PRELIMINARY; PRT; 186 AA.  
AC Q980N7;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Adenylate cyclase, cyab-type, putative (Cyab) (EC 4.6.1.1).  
GN Name=cyab; OrderedLocustNames=SS00253;  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI_TaxID=2287;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;  
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
RA Awayez M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,  
RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,  
RA Heikamp-de Jong I., Jeffries A.C., Koza C.J., Medina N., Peng X.,  
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."; Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
DR EMBL; AE006661; AAK40592.1; -: Genomic_DNA.  
DR PIR; A90167; A90167.  
DR GO; GO:0004016; F:adenylate cyclase activity; IEA.  
DR GO; GO:0016829; F:lyase activity; IEA.  
DR GO; GO:0006171; P:CAMP biosynthesis; IEA.  
DR InterPro; IPR008172; Adenylate_cyc.  
DR InterPro; IPR008173; Cyab.  
DR Pfam; PF01928; CYTH; 1.  
DR PIRSF; PIRSF005720; Ad_cyc_Cyab; 1.  
DR ProDom; PD009560; Cyab; 1.  
DR TIGRFAMs; TIGR00318; cyab; 1.  
KW Complete proteome; Lyase.  
SQ SEQUENCE 186 AA; 21820 MW; 1B47C630B438C868 CRC64;  
  
Query Match 46.4%; Score 45; DB 2; Length 186;  
Best Local Similarity 55.6%; Pred. No. 45;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 TIKGPTLRWLKSRHTS 18  
| ||| ||| ||| ||| |||  
Db 65 TYKGPKLHSLKAREIS 82  
  
RESULT 18  
Q8XR40_RALSO  
ID Q8XR40_RALSO PRELIMINARY; PRT; 243 AA.  
AC Q8XR40;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	GalU.
DE	Names:galU; OrderedLocusNames=MAP0924;
OS	Mycobacterium paratuberculosis.
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC	Mycobacterium avium complex (MAC).
OX	NCBI_TaxID=1770;
RN	[1]
RN	NUCLEOTIDE SEQUENCE.
RC	STRAIN=x10;
RL	Li L., Bermanine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL	EMBL: AS017230; AAS03241.1; -; Genomic DNA.
DR	GO: GO:0016301; P:kinase activity; IEA.
DR	GO: GO:0016775; P:nucleotidyltransferase activity; IEA.
DR	GO: GO:0016740; P:transferase activity; IEA.
DR	GO: GO:0009058; P:biosynthesis; IEA.
DR	InterPro: IPR005835; NTP transferase.
DR	Pfam: PF00483; NTP_transferase; 1.
KW	Complete proteome.
SQ	SEQUENCE 302 AA; 32149 MW; 4E5D2B1AB572BAE7 CRC64;
	Query Match 46.4%; Score 45; DB 2; Length 302;
	Best Local Similarity 72.7%; Pred. No. 78;
	Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps
Qy	4 GTTLRQWLKSR 14
	:
Db	286 GPDLRQWLVAR 296
RESULT 21	
QV126 HELHP	
ID	QV126 HELHP PRELIMINARY; PRT; 351 AA.
AC	QV126;
DT	01-OCT-2003 (TrEMBLrel. 25, Created)
DT	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Ribosomal pseudouridine synthase.
GN	OrderedLocusNames=HH0783;
OS	Helicobacter hepaticus.
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC	Helicobacteraceae; Helicobacter.
OX	NCBI_TaxID=32025;
RN	[1]
RN	NUCLEOTIDE SEQUENCE.
RC	STRAIN=ATCC 51449 / 3B1;
RC	MEDLINE=232709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA	Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA	Sell M., Droege M., Farckmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA	Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA	Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT	"The complete genome sequence of the carcinogenic bacterium
RT	Helicobacter hepaticus";
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
CC	-/- CATALYTIC ACTIVITY: rRNA uridine = rRNA pseudouridine.
DR	EMBL: AS017146; AAP7380.1; -; Genomic DNA.
DR	HSP; P33643; 1QYU.
DR	GO: GO:0016829; P:lyase activity; IEA.
DR	GO: GO:0009982; P:pseudouridine synthase activity; IEA.
DR	GO: GO:0004730; P:pseudouridylate synthase activity; IEA.
DR	GO: GO:0003723; F:RNA binding; IEA.
DR	GO: GO:0006364; P:rRNA processing; IEA.
DR	InterPro: IPR006145; Pseudou synth.
DR	InterPro: IPR006225; Pseud_synth_RLud.
DR	InterPro: IPR002942; S4.
DR	Pfam: PF00849; Pseudou_synth_2; 1.
DR	Pfam: PF01479; S4; 1.
DR	ProDom: PD001819; Pseudou_synth; 1.
DR	SMART: SM00363; S4; 1.

DR TIGRPAMS; TIGR00005; rluD_subfam; 1.
DR PROSITE; PS50889; S4; 1.
DR Complete proteome; Isomerase; RNA-binding; rRNA processing.
SQ SEQUENCE 351 AA; 40281 MW; 9C5B9C3E7733F6DE CRC64;

Query Match 46.4%; Score 45; DB 2; Length 351;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 8; Conservative 2; Mismatches 6; Indels

Qy 1 TIKPTLRQWLKSRH 16
 ::||| ||| |
p/b 114 SVKEPTLVDWKLNAH 129

RESULT 22
Q6N3E8 RHOPA
ID Q6N3E8 RHOPA PRELIMINARY; PRT; 789 AA.

AC
DT
DT
DT
DT
DE

05-JUL-2004 (TRENBLrel. 27, Created)
05-JUL-2004 (TRENBLrel. 27, Last sequence update)
05-JUL-2004 (TRENBLrel. 27, Last annotation update)
05-JUL-2004 (TRENBLrel. 27, Last annotation update)
Possible RND Superfamily transporter.

GN OrderLocusNames=KFAJ746;
 OS Rhodopseudomonas palustris.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Rhodopseudomonas.
 OX NCBI TaxID=1076;

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX Pubmed=14704707; DOI=10.1038/nbt923;

RA Larimer F.W., Chazin F., Hausner L., Mamerlin U.E., Wallinga E.P., DO
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison P.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium *Rhodospirillum rubrum* strain 12210."

RL Nat. Biotechnol. 22:55-61 (2004).
DR EMBL; BX572605; CAE29187.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR TrpasePro; IBB000731; SS0 5TM

```

DR InterPro; IPR002035; VWF A.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS50156; SSD; 1.

```

SQ	SEQUENCE	789 AA;	85214 MW;	31C1C9E9CAC530CC	CRC64;
SQ	Query Match		46.4%;	Score 45;	DB 2; Length

Matches	7; Conservative	3; Mismatches	3; Indels	0; Gaps	0;
QY	6	TLROWLKREHTS	18		

Db 523 TLRRWLSEKAHTT 535

RESULT 23
Q4QH74 LEIMA
ID Q4QH74 LEIMA PRELIMINARY; PRT; 1126 AA.

AC	Q4Q8/4	13-SEP-2005	(TrEMBLrel. 31, Created)
DT		13-SEP-2005	(TrEMBLrel. 31, Last sequence update)
DT		13-SEP-2005	(TrEMBLrel. 31, Last annotation update)

DE Protein kinase, putative (EC 2.7.1.1-).
GN ORFNames=LmjF11.0060;
OS Leishmania major.
CC Eukaryota: Euglenozoa: Kinetoplastida: Trypanosomatidae: Leishmania

OX	NCBI_TaxID=5664;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.

RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the ENML/GenBank/DBJ databases.
DR ENML: CT005250; CAJ02343.1; -; Genomic DNA.

RW	Kindberg, Translase.
SQ	SEQUENCE
	1126 AA; 124574 MW; 9F72FE6F45D18134 CRC64;
	Query Match 46.4%; Score 45; DB 2; Length 11

Q4 1 TTKGPTLRQWLKSRHTS 18
Q5 1 TTKGPTLRQWLKSRHTS 18

Db 1080 TLAETLARWIKANAATS 1097

RESULT 24
Q5TV94 ANOGA
ID Q5TV94 ANOGA PRELIMINARY; PRT; 1361 AA.

DT	01-FEB-2005	(Tremblrel. 29, Created)
DT	01-FEB-2005	(Tremblrel. 29, Last sequence update)
DT	01-FEB-2005	(Tremblrel. 29, Last annotation update)

GN ORFNAMES=ENSANG0000004323;
OS Anopheles gambiae str. PEST.
QC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

UC	Neoptera; Emdopecterygota
OC	Anophelinae; Anopheles.
OX	NCBI_TaxID=180454;
RN	[1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation".

RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP COMPLETE SEQUENCE.

RG The *Anopheles gambiae* Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

CC preliminary data.
DR EMBL; AAAB01008944; EAL41356.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.

DR	PFam; PF05041; Pecanex_C; 1.
FT	NON_TER 1 1
FT	NON_TER 1361 1361

SQ	SEQUENCE	1361 AA;	133017 MR;	43707124425310746	CR607;
	Query Match	46.4%;	Score 45;	DB 2;	Length 1361;
	Best Local Similarity	37.5%;	Pred. No. 4.1e+02;		

QY	1	TIKGPTLRQWLKSR-----EHTS	18
		.: .: .: .: .: .: .: .: .:	
Matches	9;	Conservative	5; Mismatches

Db 726 TLRSPKIMSWLQSAIEALEHTT 749

Q7PS78 ANOGA
ID Q7PS78 ANOGA PRELIMINARY; PRT; 1431 AA.
AC Q7PS78;

DT	01-MAR-2004	(TrEMBLrel. 26, Last sequence update)
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)
DE	ENSANGP00000005643	(Fragment).

OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;

```

OX NCBI_TaxID=180454;
RN (1) NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008844; EAA06040.3; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR007735; Pecanex C.
DR Pfam; PF05041; Pecanex_C; 1.
PT NON_TER 1431 1431
SQ SEQUENCE 1431 AA; 158819 MW; 5EFA1020B884C7823 CRC64;

Query Match 46.4%; Score 45; DB 2; Length 1431;
Best Local Similarity 37.5%; Pred. No. 4.4e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 6; Gaps 1;

QY 1 TIKGPTLRQWLKSR-----EHTS 18
||:| | | | |
Db 1094 TLRSPKLMGWSQAIEALEHTT 1117

RESULT 26
Q4KFC8 PSEFS
ID Q4KFC8 PSEFS PRELIMINARY; PRT; 642 AA.
AC Q4KFC8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Soluble lytic transglycosylase, putative.
GN ORFNames=PFL1936;
OS Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PF-5;
RC PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodidi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouiri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.;
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5.";
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL; CP000076; AA91223.1; -; Genomic DNA.
SQ SEQUENCE 642 AA; 73575 MW; 52DFB0524E4AD589 CRC64;

Query Match 45.9%; Score 44.5; DB 2; Length 642;
Best Local Similarity 56.4%; Pred. No. 2.2e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 4 GP-TLRQWLKSRHETS 18
||:| | | | |
Db 579 GPGRVQWLKGADHLS 594

RESULT 27
Q4JW69 CORJK
ID Q4JW69 CORJK PRELIMINARY; PRT; 131 AA.
AC Q4JW69;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)

```

```

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative ribosome-associated heat shock protein.
GN ORFNames=jk0776;
OS Corynebacterium jeikeium (strain K411).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=306537;
RN (1)
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=K411;
RC PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
RA Tauch A., Kaiser O., Hain T., Goessmann A., Weisshaar B.,
RA Albersmeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
RA Kalinowski J., Meyer F., Rupp O., Schmelker S., Viehoveer P.,
RA Puhler A.;
RT "Complete Genome Sequence and Analysis of the Multiresistant
RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requiring
RT Bacterium of the Human Skin Flora.";
RL J. Bacteriol. 187:4671-4682(2005).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=K411;
RC Linke B., Tauch A.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR931997; CA136938.1; -; Genomic DNA.
SQ Heat shock.
KW Heat shock.
SQ SEQUENCE 131 AA; 14663 MW; 59955BFC75342A18 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 131;
Best Local Similarity 46.7%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRH 16
:| | | | |
Db 57 VVGDTVRVWHHREH 71

RESULT 28
P93490 PEA
ID P93490 PEA PRELIMINARY; PRT; 161 AA.
AC P93490;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cell wall invertase II (fragment).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=var. final; TISSUE=Seed coat;
RA Buchner P.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z83339; CAB05954.1; -; mRNA.
DR PIR; T06826; T06826.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001362; Glyco_hydro_32.
DR Pfam; PF00251; Glyco_hydro_32; 1.
DR SMART; SM00640; Glyco_32; 1.
DR KW Glycosidase; Hydrolase.
FT NON_TER 1 1
FT NON_TER 161 161
SQ SEQUENCE 161 AA; 18033 MW; 32B6B07767F4ABD6 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 161;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRH 16

```

Query Match	45.4%	Score 44	DB 2	Length 302
Best Local Similarity	72.7%	Pred. No. 1.1e+02		
Matches	8	Conservative 1	Mismatches 2	Indels 0
6 TLRLWLKSRH 16				
93 TLLWLESREH 103				
RESULT 31				
P90433 SIVCZ PRELIMINARY;	PRT;	313 AA.		
P90433 SIVCZ PRELIMINARY;	PRT;	313 AA.		
01-MAY-1997 (TrEMBLrel. 03, Created)				
01-MAY-1997 (TrEMBLrel. 03, Last sequence update)				
01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
Truncated reverse transcriptase (Fragment).				
Name: pol.				
Chimpanzee immunodeficiency virus (STV(cpz)) (CIV).				
Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;				
Lentivirus; Primate lentivirus group.				
NCBI_TaxID=11723;				
NCBI_TaxID=11723;				
[1]				
NUCLEOTIDE SEQUENCE.				
Smith J.M., Kraiselburd E.N., Torres J.V.;				
Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.				
-1- SIMILARITY: Contains 1 reverse transcriptase domain.				
EMBL; U83413; AAB41428.1; -; Genomic_DNA.				
HSP; Q07387; ITCW.				
SWR; P90433; 69-167, 173-313.				
MEKPS; A02.002; -.				
GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.				
GO; GO:0008233; F:peptidase activity; IEA.				
GO; GO:0003723; F:RNA binding; IEA.				
GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.				
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.				
GO; GO:0006278; P:RNA-dependent DNA replication; IEA.				
InterPro; IPR001969; Pept Asp_A5.				
InterPro; IPR009007; Pept Aspartc_cat.				
InterPro; IPR001995; Peptidase_A2_cat.				
InterPro; IPR000477; RVTse.				
Pfam; PF000077; RVP; 1.				
Pfam; PF000078; RVT_1; 1.				
PROSITE; PS00175; ASP_PROT_RETROV; 1.				
PROSITE; PS00141; ASP_PROTEASE; 1.				
PROSITE; PS00878; RT_FOL; 1.				
Asparyl protease; Hydrolase; Protease; RNA-directed DNA polymerase.				
NON_TER				
SEQUENCE 313 AA; 34675 MW; 5A0BB016783FC8A6 CRC64;				
Query Match	45.4%	Score 44	DB 2	Length 313
Best Local Similarity	61.5%	Pred. No. 1.2e+02		
Matches	8	Conservative 2	Mismatches 3	Indels 0
3 KGPTLRLWLKSRRE 15				
184 EGPKLRQWPLSKS 196				
RESULT 32				
Q4NF14_9M1CC PRELIMINARY;	PRT;	332 AA.		
Q4NF14_9M1CC PRELIMINARY;	PRT;	332 AA.		
13-SEP-2005 (TrEMBLrel. 31, Created)				
13-SEP-2005 (TrEMBLrel. 31, Last sequence update)				
13-SEP-2005 (TrEMBLrel. 31, Last annotation update)				
DTDP-glucose 4,6-dehydratase.				
ORFNames=ArthDRAFT 1942;				
Arthrobacter sp. FB24.				
Arthrobacter sp. FB24.				
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
Micrococcales; Micrococcaceae; Arthrobacter.				
NCBI_TaxID=290399;				
[1]				

```

Best Local Similarity 37.5%; Pred.No. 1.3e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 KGPTLROWLKSREHTS 18
DB 222 KGENVRDWHTEHDSS 237
      ||| : | : : ||
      ||| : | : : ||

RESULT 34
Q53MP5_ORYSA
ID Q53MP5_ORYSA PRELIMINARY; PRT; 371 AA.
AC Q53MP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS ORFNames=LOC_Os11g19140;
OC Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Hamilton J.,
RA Jones K., Tallon L., Feldblyum T., Tsitrin T., Bera J., Kim M.,
RA Jin S., Fadrosh D., Vuong H., Overton II L., Reardon M., Weaver B.,
RA Johri S., Lewis M., Uterback T., Van Aken S., Wortman J., Haas B.,
RA Koo H., Zismann V., Haiao J., Iobst S., de Vazeilles A., White O.,
RA Salzberg S., Fraser C.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -/- PATHWAY: Lignin biosynthesis.
DR EMBL; AC135599; AX96798.1; -; Genomic DNA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR InterPro; IPR000051; SAM_bd.
KW Hypothetical protein; Lignin biosynthesis; Methyltransferase;
KW Transferase.
SQ SEQUENCE 371 AA; 41039 MW; 78C040F7BF64961F CRC64;

Query Match 45.4%; Score 44; DB 2; Length 371;
Best Local Similarity 69.2%; Pred.No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 GPTLROWLKSREH 16
DB 131 GETRESLKRREH 143
      ||||| : ||||
      ||||| : ||||

RESULT 35
Q82PX5_STRAW
ID Q82PX5_STRAW PRELIMINARY; PRT; 377 AA.
AC Q82PX5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS OrderedLocusNames=SAV747;
OC Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
```

Nat. Biotechnol. 21:526-531(2003).

[2]
[2] NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,
Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "genome sequence of an industrial microorganism Streptomyces
RT aerogenus deducing the ability of producing secondary
metabolites.";
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
EMBL; BA000300; BAC68457.1; -, Genomic_DNA.
Complete proteome; Hypothetical protein.
SEQUENCE 377 AA; 41307 MW; 0253176AAABE62F3 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 377;
Best Local Similarity 46.7%; Pred No. 1.5e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IKGPTFLQWLKSREH 16
 ::||| |||| |:
DB 168 MEGPDLRWLPNRY 182

RESULT 36
Q4PED7 USTMA PRELIMINARY; PRT; 472 AA.

ID Q4PED7 USTMA PRELIMINARY;
AC Q4PED7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DI 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DR 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNAMES=UM01526.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
NCBI_TaxID=237631;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521.
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
Bayul T., Blitshteyn B., Bloom T., Bye J., Boguclavsky L.,
Rowewy M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
CA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
Drjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
Edrickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
Gnrke A., Guyette A., Graham J., Grandbois E., Gyaltzen K., Hafez N.,
Haegopian D., Hagos B., Hall J., Hatchler B., Heller A., Higgins H.,
Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
Jaife D., Jones C., Kamal M., Katam A., Kamysselellis M., Karlsson E.,
Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
LA Lindblad-Toh K., Liu X., Lokytang T., Lokytang Y., Lucien O.,
Lui A., Ma L.J., Mabbutt R., Macdonald J., Maclean C., Major J.,
Manning J., Marabellia R., Maru K., Matthews C., Maucelli E.,
McCarthy M., McDonough S., Mcghee T., Meldrum J., Menaus L.,
ME Meirov J., Mihailov A., Mihova T., Mikkelson T., Mieneus V., Moru K.,
Mozer J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
NGuyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'Leary S., Omotosho B.,
O'Neill K., Oman S., Parker S., Perrin D., Phunkhang P., Pigani B.,
Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
Rutman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougneiz C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stefgon K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MPS membrane transporter (Fragment).
GN Names=mt;
OS Gibberella fujiuroi (Bakanae and foot rot disease fungus) (Fusarium moniliforme).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella;
OC Gibberella fujiuroi complex.
OX NCBI_TaxID=5127;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF293;
RX MEDLINE=21416226; PubMed=11525413; DOI=10.1007/s002940100218;
RA Voss T., Schulte J., Tudzynski B.;
RT "A new MPS transporter gene next to the gibberellin biosynthesis gene cluster of Gibberella fujiuroi is not involved in gibberellin secretion."
RL Curr. Genet. 39:377-383(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AJ272424; CAB75959.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Transport.
FT NON_TER 1
SQ SEQUENCE 648 AA; 72250 MW; 4C90EEB49E25AF9C CRC64;
Query Match 45.4%; Score 44; DB 2; Length 648;
Best Local Similarity 77.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 9 QWLKSRHT 17
Db 122 QWLKSKHT 130
RESULT 39
Q4WKMO ASPFU PRELIMINARY; PRT; 764 AA.
AC Q4WKMO;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE DNA mismatch repair protein Msh4, putative.
GN ORFNames=Afulg02000;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF293;
RA Niernan W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Bertram M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Laiton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mounya I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Penalva M.A., Perle M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus".
RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
CC -!- SIMILARITY: Belongs to the DNA mismatch repair mutS family.
DR EMBL; RAHF01000007; EAL87912.1; -; Genomic_DNA.
DR InterPro; IPR000432; MutS_C.
DR InterPro; IPR007860; MutS_II.
DR InterPro; IPR007696; MutS_III.
DR InterPro; IPR007861; MutS_IV.
DR Pfam; PF05188; MutS_II; 1.
DR Pfam; PF05192; MutS_III; 1.
DR Pfam; PF05190; MutS_IV; 1.
DR Pfam; PF00488; MutS_V; 1.
DR ProDom; PD001263; MutS_C; 1.
DR SMART; SM00534; MUTSac; 1.
DR SMART; SM00533; MUTSc; 1.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR DNA-binding.
KW DNA-binding.
SQ SEQUENCE 764 AA; 84832 MW; 8870792DB18E5514 CRC64;
Query Match 45.4%; Score 44; DB 2; Length 764;
Best Local Similarity 58.3%; Pred. No. 3.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLK 12
Db 707 TLEGDALREWLK 718
RESULT 40
Q6VG40 SIVCZ PRELIMINARY; PRT; 1017 AA.
AC Q6VG40;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Pol protein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22972551; PubMed=14610175;
RX DOI=10.1128/JVI.77.23.12523-12534.2003;
RA Courgnaud V., Abela B., Pourrut X., Mpoudi-Ngole E., Loul S.,
RA Delaporte E., Peeters M.;
RT "Identification of a new simian immunodeficiency virus lineage with a vpu gene present among different cercopithecus monkeys (C. mona, C. cephus, and C. nictitans) from Cameroon.";
RL J. Virol. 77:12523-12534(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts the viral RNA genome into double-stranded viral DNA. The RNase H domain of the reverse transcriptase performs two functions. It degrades the RNA template and specifically removes the RNA primer from the RNA/DNA hybrid. Following nuclear import, the integrase catalyzes the insertion of the linear, double-stranded viral DNA into the host cell chromosome. Endogenous Pol proteins may have kept, lost or modified their original function during evolution (By similarity).
CC -!- DOMAIN: The LTRG and YXDD motifs are catalytically important and conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY340701; AAR02377.1; -; Genomic_DNA.


```
Best Local Similarity 61.5%; Pred. No. 4.4e+02; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3;

QY 3 KGPTLRQWLKSR 15
   :||| ||||| :|
Db 184 EGPQLRQWPLSK 196

RESULT 42
P89154 SIVCZ
ID P89154 SIVCZ PRELIMINARY; PRT; 1019 AA.
AC P89154;
DT 01-MAY-1997 (T-REMBLrel. 03, Created)
DT 13-SEP-2005 (T-REMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-REMBLrel. 31, Last annotation update)
DE Pol protein (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SiVemE543;
RX MEDLINE=97151152; PubMed=8995688;
RA Hirsch V., Adger-Johnson D., Campbell B., Goldstein S., Brown C.,
RA Elkins W.R., Montefiori D.C.;
RT "A molecularly cloned, pathogenic, neutralization-resistant simian
RT immunodeficiency virus, SiVemE543-3.";
RL J. Virol. 71:1608-1620(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SiVemE543;
RA Ourmanov I.K., Deighani H., Kuwata T., Hirsch V.M.;
RA Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; U72748; AAC56559.2; -; Genomic_DNA.
DR SMR; P89154; 69-167, 781-939.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept_Asp_AS.
DR InterPro; IPR009007; Pept_Aspartic_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT_connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; Rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT_connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
```

```
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNase_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115586 MW; 2AB14CABAC66FF00 CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLRQWLKSR 15
   :||| ||||| :|
Db 184 EGPQLRQWPLSK 196

RESULT 43
Q7ZBR7 SIVCZ
ID Q7ZBR7 SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR7;
DT 01-JUN-2003 (T-REMBLrel. 24, Created)
DT 01-JUN-2003 (T-REMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-REMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name=pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OX NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RA DOI=10.1126/JVI.77.11.6405-6418.2003;
RA Deighani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5.";
J. Virol. 77:6405-6418(2003).
CC -!- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- DOMAIN: The LPQG and YXDD motifs are catalytically important and
CC conserved among many retroviruses (By similarity).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AX221514; AAO67307.1; -; Genomic_DNA.
DR HSSP; P04584; 1M02.
DR SMR; Q7ZBR7; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; F:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
```

```
DR InterPro; IPR001969; Pept Asp AS.
DR InterPro; IPR009007; Pept Aspartic cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
DR Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115341 MW; A886525DF1B26F CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLROWLSRE 15
Db 184 EGPRLROWPLSKE 196
:::|||||:|

RESULT 44
Q7ZBR5 SIVCZ
ID Q7ZBR5 SIVCZ PRELIMINARY; PRT; 1019 AA.
AC Q7ZBR5; 2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pol (Fragment).
GN Name-pol;
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
OC Lentivirus; Primate lentivirus group.
OC NCBI_TaxID=11723;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22628501; PubMed=12743298;
RX DOI=10.1128/JVI.77.11.6405-6418.2003;
RA Dehghani H., Puffer B.A., Doms R.W., Hirsch V.M.;
RT "Unique pattern of convergent envelope evolution in simian
RT immunodeficiency virus-infected rapid progressor macaques: association
RT with CD4-independent usage of CCR5;"
RL J. Virol. 77:6405-6418(2003).
CC -1- FUNCTION: Early post-infection, the reverse transcriptase converts
CC the viral RNA genome into double-stranded viral DNA. The RNase H
CC domain of the reverse transcriptase performs two functions. It
CC degrades the RNA template and specifically removes the RNA primer
CC from the RNA/DNA hybrid. Following nuclear import, the integrase
CC catalyzes the insertion of the linear, double-stranded viral DNA
CC into the host cell chromosome. Endogenous Pol proteins may have
CC kept, lost or modified their original function during evolution
CC (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: the LPQG and YXDD motifs are catalytically important and
```

```
CC conserved among many retroviruses (By similarity).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; AY221515; AAO67309.1; -; Genomic_DNA.
DR HSSP; P04584; 1MU2.
DR SMR; Q7ZBR5; 69-167, 781-939.
DR MEROPS; A02.002; -.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008907; F:integrase activity; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0015074; P:DNA integration; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_N_Zn_bd.
DR InterPro; IPR001969; Pept Asp AS.
DR InterPro; IPR009007; Pept_Aspartc_cat.
DR InterPro; IPR001995; Peptidase_A2_cat.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR012337; RNaseH_fold.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR010659; RVT connect.
DR InterPro; IPR010661; RVT_thumb.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00552; Integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; RNaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00077; RVP; 1.
DR Pfam; PF00078; RVT_1; 1.
DR Pfam; PF06815; RVT connect; 1.
DR Pfam; PF06817; RVT_thumb; 1.
DR PROSITE; PS01175; ASP_PROT_RETROV; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS50994; INTEGRASE; 1.
DR PROSITE; PS51027; INTEGRASE_DBD; 1.
DR PROSITE; PS50879; RNASE_H; 1.
DR PROSITE; PS50878; RT_POL; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
DR PROSITE; PS50876; ZF_INTEGRASE; 1.
KW Aspartyl protease; DNA integration; Hydrolase; Multifunctional enzyme;
KW Protease; RNA-directed DNA polymerase; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1019 AA; 115613 MW; 6002D54F14648CBC CRC64;

Query Match 45.4%; Score 44; DB 2; Length 1019;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 KGPTLROWLSRE 15
Db 184 EGPRLROWPLSKE 196
:::|||||:|

RESULT 45
POL_HV2D2
ID POL_HV2D2 STANDARD; PRT; 1058 AA.
AC P15833;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Pol polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.47);
DE Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (RT);
DE Integrase (IN)].
GN Name=POL;
OS Human immunodeficiency virus type 2 (isolate D205.7) (HIV-2).
```

OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;
 OC Lentivirus; Primate lentivirus group.
 OX NCBI_TaxID=11716;
 RN [1]
 RP MEDLINE=9008181; PubMed=2594088; DOI=10.1038/342948a0;
 RA Dietrich U., Adamaki M., Kreutz R., Seipp A., Kuehnel H.,
 RA Ruebsamen-Waigmann H.;
 RT "A highly divergent HIV-2-related isolate.";
 RL Nature 342:948-950(1989).
 CC -1- FUNCTION: During replicative cycle of retroviruses, the reverse-
 transcribed viral DNA is integrated into the host chromosome by
 CC the viral integrase enzyme. RNase H activity is associated with
 CC the reverse transcriptase.
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 phosphomonoester.
 CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
 CC diphosphate + DNA(n+1).
 CC -1- CATALYTIC ACTIVITY: Endopeptidase for which the P1 residue is
 CC preferably hydrophobic.
 CC -1- PTM: Cleavage sites that yield the mature proteins remain to be
 CC determined.
 CC -1- SIMILARITY: Belongs to the retroviruses Pol polyprotein family.
 CC -1- SIMILARITY: Contains 1 integrase catalytic domain.
 CC -1- SIMILARITY: Contains 1 integrase-type DNA-binding domain.
 CC -1- SIMILARITY: Contains 1 integrase-type zinc finger.
 CC -1- SIMILARITY: Contains 1 peptidase A2 domain.
 CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
 CC -1- SIMILARITY: Contains 1 RNase H domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; X61240; -; NOT_ANNOTATED_CDS; Genomic_DNA.
 DR PIR; S08436; S08436.
 DR HSPR; P04584; 1MU2.
 DR SRR; P15833; 106-204, 207-758, 817-974.
 DR HTV; X16109; POLS2D205.
 DR InterPro; IPR001037; Integrase_C.
 DR InterPro; IPR003308; Integrase_N_zn_bd.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR InterPro; IPR009007; Pept_Asp_cat.
 DR InterPro; IPR001995; Peptidase_A2_cat.
 DR InterPro; IPR002156; RNaseH.
 DR InterPro; IPR012337; RNaseH_fold.
 DR InterPro; IPR001584; Rve.
 DR InterPro; IPR010659; RVT_connect.
 DR InterPro; IPR010661; RVT_thumb.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF00552; Integrase; 1.
 DR Pfam; PF02022; Integrase_zn; 1.
 DR Pfam; PF00075; RNaseH; 1.
 DR Pfam; PF00685; rve; 1.
 DR Pfam; PF00077; RVP; 1.
 DR Pfam; PF00078; RVT; 1.
 DR Pfam; PF06815; RVT_connect; 1.
 DR Pfam; PF06817; RVT_thumb; 1.
 DR PROSITE; PS50175; ASP_PROT_RETROV; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 DR PROSITE; PS50994; INTEGRASE; 1.
 DR PROSITE; PS51027; INTEGRASE_DBD; 1.
 DR PROSITE; PS50879; RNASE_H; 1.
 DR PROSITE; PS50878; RT_POL; 1.
 DR PROSITE; PS50876; ZF_INTEGRASE; 1.
 KW AIDS; Aspartyl protease; DNA integration; DNA recombination;
 KW Endonuclease; Hydrolase; Metal-binding; Multifunctional enzyme;
 KW Nuclease; Nucleotidyltransferase; Polypeptide; Protease;
 KW RNA-directed DNA polymerase; Transferase; Zinc; Zinc-finger.
 FT CHAIN 106 204 Peptase.
 FT DOMAIN 125 194 Peptidase A2.

FT DOMAIN 248 437 Reverse transcriptase.
 FT DOMAIN 636 759 RNase H.
 FT DOMAIN 815 966 Integrase catalytic.
 FT ZN_FING 765 806 Integrase-type.
 FT DNA_BIND 985 1032 Integrase-type.
 FT ACT_SITE 130 130 By similarity.
 SQ SEQUENCE 1058 AA; 119964 MW; 914D5433694B57F4 CRC64;
 Query Match 45.4%; Score 44; DB 1; Length 1058;
 Best Local Similarity 66.7%; Pred. No. 4.6e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GPTLRQWLKRSRE 15
 Db 222 GPKIRQWPLSR 233
 RESULT 46
 OSBFN8_EWENI PRELIMINARY; PRT; 1387 AA.
 ID OSBFN8_EWENI PRELIMINARY; PRT; 1387 AA.
 AC OSBFN8;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AN0642.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=227321;
 RN [1]_
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nusbbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Genome Sequence of Aspergillus nidulans";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AACD01000008; EAA65185.1; -; Genomic_DNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0005786; C:signal recognition particle (sensu Eukaryota); IEA.
 DR GO; GO:0008312; F:7S RNA binding; IEA.
 DR GO; GO:0048872; F:metal ion binding; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006605; P:protein targeting; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
 KW Transcription; Transcription regulation; Zinc.
 SQ SEQUENCE 1387 AA; 152683 MW; 0F30A6D4E00742BD CRC64;
 Query Match 45.4%; Score 44; DB 2; Length 1387;
 Best Local Similarity 44.4%; Pred. No. 6.2e+02;

QY 3 KGP---TLRQWLKREHT 17
||| ||||:|:
Db 470 KGPYAEQLRQWVKTQRET 487

ID	Q88L07_PSEPK	PRT;	657 AA.
AC	Q88L07;		
DT	01-JUN-2003 (TrEMBLrel. 24, Created)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Soluble lytic transglycosylase, putative.		
OS	OrderediocusNames=PP2130;		
GN	Pseudomonas putida (strain KT2440).		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;		
OC	Pseudomonadaceae; Pseudomonas.		
NCBI_TaxID=160488;			
[1]			
NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=22423060; PubMed=12534463;		
RX	DOI=10.1046/j.1462-2920-2002.00366.x;		
RX	Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,		
RA	Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,		
RA	Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,		
RA	Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,		
RA	Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,		
RA	Moazez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,		
RA	Wedler H., Lauber J., Stjepandic D., Hohseisel J., Straetz M., Heim S.,		
RA	Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuemmler B.,		
RA	Fraser C.M.;		
RA	"Complete genome sequence and comparative analysis of the		
RL	metabolically versatile Pseudomonas putida KT2440.";		
RL	Environ. Microbiol. 4:799-808(2002).		
DR	EMBL; AE016782; AAN67743.1; -; Genomic_DNA.		

DR TIGR; PP2130; -.
DR GO; GO:0042597; C:periplasmic space; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR InterPro; IPR008258; LT catalytic.
DR InterPro; IPR008939; LT_superhelical_U.
DR Pfam; PF01484; SLT; 1.
DR Complete proteome.
KW 03402005CD012200 CBC64.

```

Query Match          44.8%; Score 43.5; DB 2; Length 657;
Best Local Similarity 56.2%; Pred. No. 3.2e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

      4 GP-TLRWLKSRHSTS 18
      ||| :||| :|||
Db    594 GPGVRQWLKGAKHLS 609

RESULT 50
Q98183 MCV1
ID Q98183 MCV1 PRELIMINARY; PRT; 63 AA.
AC Q98183_012598; 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE MC012L (Hypothetical protein B-M,N,L.2).
GN Name=MC012L; Synonyms=B-M,N,L.2;
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OC NCBI_taxid=10280;
CN [1]
RN

```

RX MEDLINE=96325459; PubMed=8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Datarai G.,
RA Mosses B.;

RT "Genome sequence of a human tumorigenic poxvirus: prediction of
 RT specific host response-evasion genes.";
 RL Science 273:813-816(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97093414; PubMed=8938976;
 RA Martin-Gallardo A., Moratilla M., Funes J.M., Agromayor M., Nunez A.,
 RA Varas A.J., Collado M., Valencia A., Lopez-Esteban J.L.,
 RA Esteban M.;
 RT "Sequence analysis of a Molluscum contagiosum virus DNA region which
 RT includes the gene encoding protein kinase 2 and other genes with
 RT unique organization";
 RL Virus Genes 13:19-29(1996).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97352177; PubMed=9208457; DOI=10.1023/A:1007991508159;
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,
 RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;
 RT "A random DNA sequencing, computer-based approach for the generation
 RT of a gene map of Molluscum contagiosum virus.";
 RL Virus Genes 14:73-80(1997).
 DR EMBL; U60315; AAC55140.1; -; Genomic_DNA.
 DR EMBL; U66894; AAB57932.1; -; Genomic_DNA.
 DR PIR; T30614; T30614.
 KW Hypothetical protein.
 SQ SEQUENCE 63 AA; 7088 MW; 1C96B36D3E5D8F27 CRC64;

Query Match 44.3%; Score 43; DB 2; Length 63;
 Best Local Similarity 41.2%; Pred. No. 29;
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 IKGPTLRQWLKSRHTS 18
 : ||||| : : :
 Db 41 VLGETLRTWSRSKRNTA 57

Search completed: May 12, 2006, 10:49:45
 Job time : 163.385 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2006, 10:19:39 ; Search time 130.769 seconds

(without alignments)
60.479 Million cell updates/sec

Title: US-10-632-388-80

Perfect score: 97

Sequence: 1 TIKGPTLRQLKSHREHTS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A_Geneseq_21.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	18	2	Aaw09499 Thrombopo
2	97	100.0	18	2	Aaw09459 Thrombopo
3	97	100.0	18	2	Aaw36650 Thrombopo
4	97	100.0	18	2	Aaw33026 Thrombopo
5	97	100.0	18	3	Aab17024 TPO-mimet
6	97	100.0	18	4	Aau25869 Human thr
7	97	100.0	18	4	Aau25823 Human thr
8	97	100.0	18	5	Aab72910 TPO mimet
9	97	100.0	18	8	Adj73062 TPO mimet
10	97	100.0	18	8	Adj52697 CH1 delet
11	97	100.0	18	8	Adj51658 CH1 delet
12	63	64.9	19	2	Aaw09491 Thrombopo
13	63	64.9	19	2	Aaw35418 Thrombopo
14	63	64.9	19	2	Aaw36642 Thrombopo
15	63	64.9	19	4	Aau25861 Human thr
16	63	64.9	19	4	Aau25998 Human thr
17	62	63.9	18	5	Abp51693 TPO mimet
18	62	63.9	18	5	Abp51691 TPO mimet
19	62	63.9	18	8	Adq16625 TPO mimet
20	62	63.9	18	8	Adq16629 TPO mimet
21	62	63.9	18	9	Adv44363 Agonist T
22	62	63.9	18	9	Adv44367 Agonist T
23	62	63.9	18	9	Aeb12840 TPO mimet
24	62	63.9	18	9	Aeb12836 TPO mimet

25	62	63.9	19	2	Aaw09493	Aaw09493 Thrombopo
26	62	63.9	19	2	Aaw36644	Aaw36644 Thrombopo
27	62	63.9	19	4	Aau25863	Aau25863 Human thr
28	62	63.9	144	6	ABG71748	ABG71748 Antibody
29	58	59.8	18	4	Aau26006	Aau26006 Human thr
30	58	59.8	18	2	Aaw09460	Aaw09460 Thrombopo
31	58	59.8	18	2	Aaw09498	Aaw09498 Thrombopo
32	58	59.8	18	2	Aaw36649	Aaw36649 Thrombopo
33	58	59.8	18	2	Aaw33027	Aaw33027 Thrombopo
34	58	59.8	18	2	Aaw36652	Aaw36652 Thrombopo
35	58	59.8	18	3	AAB17026	AAB17026 TPO-mimet
36	58	59.8	18	4	Aau25868	Aau25868 Human thr
37	58	59.8	18	4	Aau25824	Aau25824 Human thr
38	58	59.8	18	4	Aau25871	Aau25871 Human thr
39	58	59.8	18	5	ABB72912	ABB72912 TPO mimet
40	58	59.8	18	7	ADJ73064	ADJ73064 TPO mimet
41	58	59.8	18	8	ADJ52699	ADJ52699 CH1 delet
42	58	59.8	18	8	ADJ51660	ADJ51660 CH1 delet
43	58	59.8	18	9	ADV44468	ADV44468 Agonist T
44	58	59.8	18	9	AEb12939	AEb12939 TPO mimet
45	58	59.8	128	8	ADQ16705	ADQ16705 Modified
46	58	59.8	128	9	ADV44466	ADV44466 Anti-teta
47	58	59.8	128	9	ADV44463	ADV44463 Anti-teta
48	58	59.8	128	9	ADV44467	ADV44467 Anti-teta
49	58	59.8	128	9	ADV44465	ADV44465 Anti-teta
50	58	59.8	128	9	ADV44441	ADV44441 pAX116 va
51	58	59.8	128	9	ADV44464	ADV44464 Anti-teta
52	58	59.8	128	9	AEb12934	AEb12934 Antibody
53	58	59.8	128	9	AEb12935	AEb12935 Antibody
54	58	59.8	128	9	AEb12913	AEb12913 Antibody
55	58	59.8	128	9	AEb12936	AEb12936 Antibody
56	58	59.8	128	9	AEb12937	AEb12937 Antibody
57	58	59.8	128	9	AEb12938	AEb12938 Antibody
58	58	59.8	132	9	ADV44473	ADV44473 Anti-teta
59	58	59.8	150	9	AEb12945	AEb12945 Antibody
60	58	59.8	225	8	ADQ16704	ADQ16704 Modified
61	58	59.8	234	9	AEb12912	AEb12912 Antibody
62	58	59.8	249	9	ADV44440	ADV44440 pAX116 va
63	58	59.8	282	9	AEb12930	AEb12930 Antibody
64	58	59.8	459	9	ADV44459	ADV44459 Anti-teta
65	57	58.8	13	2	AAW36779	AAW36779 Thrombopo
66	57	58.8	14	2	AAW09463	AAW09463 Thrombopo
67	57	58.8	14	2	AAW09468	AAW09468 Thrombopo
68	57	58.8	14	2	AAW33030	AAW33030 Thrombopo
69	57	58.8	14	2	AAW33034	AAW33034 Thrombopo
70	57	58.8	14	2	AAW36774	AAW36774 Thrombopo
71	57	58.8	14	2	ADI24843	ADI24843 AF 12505
72	57	58.8	14	3	AAy96515	AAy96515 Thrombopo
73	57	58.8	14	3	AAAB16962	AAAB16962 TPO-mimet
74	57	58.8	14	4	AAU25827	AAU25827 Human thr
75	57	58.8	14	4	AAU26037	AAU26037 Human thr
76	57	58.8	14	4	AAU26004	AAU26004 Human thr
77	57	58.8	14	4	ABT72853	ABT72853 TPO mimet
78	57	58.8	14	5	ABP51669	ABP51669 Thrombopo
79	57	58.8	14	5	AAE18011	AAE18011 Human lig
80	57	58.8	14	6	ABG71747	ABG71747 TPO recep
81	57	58.8	14	7	ABR62907	ABR62907 Thrombopo
82	57	58.8	14	7	ADN33697	ADN33697 Erythropo
83	57	58.8	14	7	ADN59652	ADN59652 Thrombopo
84	57	58.8	14	8	ADL27293	ADL27293 Amino aci
85	57	58.8	14	8	ADM72503	ADM72503 TPO mimet
86	57	58.8	14	8	ADM72483	ADM72483 TPO mimet
87	57	58.8	14	8	ADM72526	ADM72526 TPO mimet
88	57	58.8	14	8	ADM72487	ADM72487 TPO mimet
89	57	58.8	14	8	ADQ16584	ADQ16584 Agonist T
90	57	58.8	14	8	ADT92482	ADT92482 Thrombopo
91	57	58.8	14	9	ADU70210	ADU70210 Thrombopo
92	57	58.8	14	9	ADU70206	ADU70206 Thrombopo
93	57	58.8	14	9	ADU75982	ADU75982 Peptide-b
94	57	58.8	14	9	ADU75979	ADU75979 Thrombopo
95	57	58.8	14	9	ADV44319	ADV44319 Agonist T
96	57	58.8	14	9	AEb12792	AEb12792 TPO mimet
97	57	58.8	15	2	AAW35416	AAW35416 Thrombopo

98	57	58.8	15	2	AAW36780	Aaw36780	Thrombopo	171	57	58.8	18	8	ADQ16605	Adq16605	TPO mimet
99	57	58.8	15	2	AAW36776	Aaw36776	Thrombopo	172	57	58.8	18	8	ADQ16609	Adq16609	TPO mimet
100	57	58.8	15	2	AAW66714	Aaw66714	Peptide c	173	57	58.8	18	9	ADV44335	Adv44335	Agonist T
101	57	58.8	15	2	AAW66721	Aaw66721	Peptide c	174	57	58.8	18	9	ADV44345	Adv44345	Agonist T
102	57	58.8	15	2	AAW66712	Aaw66712	Peptide c	175	57	58.8	18	9	ADV44351	Adv44351	Agonist T
103	57	58.8	15	3	AAW20684	Aaw20684	Thrombocy	176	57	58.8	18	9	ADV44357	Adv44357	Agonist T
104	57	58.8	15	4	AAU25996	Aau25996	Human thr	177	57	58.8	18	9	ADV44349	Adv44349	Agonist T
105	57	58.8	15	4	AAU26026	Aau26026	Human thr	178	57	58.8	18	9	ADV44343	Adv44343	Agonist T
106	57	58.8	15	4	AAU26020	Aau26020	Human thr	179	57	58.8	18	9	ADV44347	Adv44347	Agonist T
107	57	58.8	15	4	AAU25831	Aau25831	Human thr	180	57	58.8	18	9	ADV44384	Adv44384	Agonist T
108	57	58.8	15	4	AAU26007	Aau26007	Human thr	181	57	58.8	18	9	ADV44353	Adv44353	Agonist T
109	57	58.8	15	5	ABP51670	Abp51670	Thrombopo	182	57	58.8	18	9	ADV44361	Adv44361	Agonist T
110	57	58.8	15	7	ABR62908	AbR62908	Thrombopo	183	57	58.8	18	9	ADV44379	Adv44379	Agonist T
111	57	58.8	15	8	ADM72485	Adm72485	TPO mimet	184	57	58.8	18	9	ADV44344	Adv44344	Agonist T
112	57	58.8	15	8	ADM72479	Adm72479	TPO mimet	185	57	58.8	18	9	ADV44359	Adv44359	Agonist T
113	57	58.8	15	8	ADM72502	Adm72502	TPO mimet	186	57	58.8	18	9	ADV44365	Adv44365	Agonist T
114	57	58.8	15	8	ADM72492	Adm72492	TPO mimet	187	57	58.8	18	9	AEb12826	Eb12826	TPO mimet
115	57	58.8	15	8	ADM72478	Adm72478	TPO mimet	188	57	58.8	18	9	AEb12820	Eb12820	TPO mimet
116	57	58.8	15	8	ADM72533	Adm72533	TPO mimet	189	57	58.8	18	9	AEb12857	Eb12857	Antibody
117	57	58.8	15	8	ADM72490	Adm72490	TPO mimet	190	57	58.8	18	9	AEb12834	Eb12834	TPO mimet
118	57	58.8	15	8	ADM72486	Adm72486	TPO mimet	191	57	58.8	18	9	AEb12830	Eb12830	TPO mimet
119	57	58.8	15	8	ADM72491	Adm72491	TPO mimet	192	57	58.8	18	9	AEb12832	Eb12832	TPO mimet
120	57	58.8	15	8	ADM72522	Adm72522	TPO mimet	193	57	58.8	18	9	AEb12852	Eb12852	TPO mimet
121	57	58.8	15	8	ADM72523	Adm72523	TPO mimet	194	57	58.8	18	9	AEb12816	Eb12816	TPO mimet
122	57	58.8	15	8	ADM72493	Adm72493	TPO mimet	195	57	58.8	18	9	AEb12818	Eb12818	TPO mimet
123	57	58.8	15	8	ADM72482	Adm72482	TPO mimet	196	57	58.8	18	9	AEb12828	Eb12828	TPO mimet
124	57	58.8	15	8	ADQ16585	Adq16585	TPO mimet	197	57	58.8	18	9	AEb12822	Eb12822	TPO mimet
125	57	58.8	15	8	ADT92483	Adt92483	Modified	198	57	58.8	18	9	AEb12824	Eb12824	TPO mimet
126	57	58.8	15	9	ADU70209	Adu70209	Thrombopo	199	57	58.8	18	9	AEb12838	Eb12838	TPO mimet
127	57	58.8	15	9	ADU75981	Adu75981	Peptide-b	200	57	58.8	19	5	ABb73391	Bb73391	TPO-mimet
128	57	58.8	15	9	ADV44320	Adv44320	Agonist T	201	57	58.8	19	5	ABb73390	Bb73390	TPO-mimet
129	57	58.8	15	9	ABE12793	AbE12793	TPO mimet	202	57	58.8	20	3	AAb18003	Ab18003	FC-TMP pe
130	57	58.8	16	2	AAW19534	Aaw19534	Thrombopo	203	57	58.8	20	3	AAb17929	Ab17929	TPO-mimet
131	57	58.8	16	2	AAW33035	Aaw33035	Thrombopo	204	57	58.8	20	5	ABb73403	Bb73403	TPO mimet
132	57	58.8	16	2	AAW36775	Aaw36775	Thrombopo	205	57	58.8	21	7	ADN59687	Adn59687	Thrombopo
133	57	58.8	16	2	AAW36771	Aaw36771	Thrombopo	206	57	58.8	22	7	ADN59819	Adn59819	TMP pepti
134	57	58.8	16	2	AAW66709	Aaw66709	Peptide c	207	57	58.8	22	8	ADQ16714	Adq16714	Immunoglo
135	57	58.8	16	2	AAW66713	Aaw66713	Peptide c	208	57	58.8	22	8	ADQ16713	Adq16713	Immunoglo
136	57	58.8	16	2	AAW66733	Aaw66733	Peptide c	209	57	58.8	22	8	ADQ16709	Adq16709	Immunoglo
137	57	58.8	16	2	AAW66716	Aaw66716	Peptide c	210	57	58.8	22	8	ADQ16706	Adq16706	Immunoglo
138	57	58.8	16	4	AAU26021	Aau26021	Human thr	211	57	58.8	22	8	ADQ16699	Adq16699	TPO mimet
139	57	58.8	16	4	AAU26005	Aau26005	Human thr	212	57	58.8	22	8	ADQ16712	Adq16712	Immunoglo
140	57	58.8	16	4	AAU26043	Aau26043	Human thr	213	57	58.8	22	8	ADQ16707	Adq16707	Immunoglo
141	57	58.8	16	4	AAU25832	Aau25832	Human thr	214	57	58.8	22	8	ADQ16711	Adq16711	Immunoglo
142	57	58.8	16	8	ADM72532	Adm72532	TPO mimet	215	57	58.8	22	8	ADQ16708	Adq16708	Immunoglo
143	57	58.8	16	8	ADM72484	Adm72484	TPO mimet	216	57	58.8	22	8	ADQ16710	Adq16710	Immunoglo
144	57	58.8	18	3	ABP16957	Abp16957	PEGylated	217	57	58.8	22	9	ADV44435	Adv44435	Modified
145	57	58.8	18	3	ABP16956	Abp16956	PEGylated	218	57	58.8	22	9	ADV44449	Adv44449	Anti-teta
146	57	58.8	18	5	ABP51687	Abp51687	TPO mimet	219	57	58.8	22	9	ADV44443	Adv44443	Anti-teta
147	57	58.8	18	5	ABP51689	Abp51689	TPO mimet	220	57	58.8	22	9	ADV44444	Adv44444	Anti-teta
148	57	58.8	18	5	ABP51688	Abp51688	TPO mimet	221	57	58.8	22	9	ADV44448	Adv44448	Anti-teta
149	57	58.8	18	5	ABP51677	Abp51677	TPO mimet	222	57	58.8	22	9	ADV44442	Adv44442	Anti-teta
150	57	58.8	18	5	ABP51686	Abp51686	TPO mimet	223	57	58.8	22	9	ADV44447	Adv44447	Anti-teta
151	57	58.8	18	5	ABP51674	Abp51674	TPO mimet	224	57	58.8	22	9	ADV44445	Adv44445	Anti-teta
152	57	58.8	18	5	ABP51684	Abp51684	TPO mimet	225	57	58.8	22	9	ADV44450	Adv44450	Anti-teta
153	57	58.8	18	5	ABP51683	Abp51683	TPO mimet	226	57	58.8	22	9	ADV44446	Adv44446	Anti-teta
154	57	58.8	18	5	ABP51685	Abp51685	TPO mimet	227	57	58.8	22	9	AEb12918	Eb12918	TPO mimet
155	57	58.8	18	5	ABP51673	Abp51673	TPO mimet	228	57	58.8	22	9	AEb12917	Eb12917	TPO mimet
156	57	58.8	18	5	ABP51690	Abp51690	TPO mimet	229	57	58.8	22	9	AEb12919	Eb12919	TPO mimet
157	57	58.8	18	5	ABP51675	Abp51675	TPO mimet	230	57	58.8	22	9	AEb12986	Eb12986	TT antibo
158	57	58.8	18	5	ABP51692	Abp51692	TPO mimet	231	57	58.8	22	9	AEb12920	Eb12920	TPO mimet
159	57	58.8	18	7	ADN59812	Adn59812	Thrombopo	232	57	58.8	22	9	AEb12916	Eb12916	TPO mimet
160	57	58.8	18	8	ADQ16661	Adq16661	TPO mimet	233	57	58.8	22	9	AEb12914	Eb12914	TPO mimet
161	57	58.8	18	8	ADQ16619	Adq16619	TPO mimet	234	57	58.8	22	9	AEb12921	Eb12921	TPO mimet
162	57	58.8	18	8	ADQ16662	Adq16662	TPO mimet	235	57	58.8	22	9	AEb12907	Eb12907	TT antibo
163	57	58.8	18	8	ADQ16641	Adq16641	TPO mimet	236	57	58.8	22	9	AEb12915	Eb12915	TPO mimet
164	57	58.8	18	8	ADQ16646	Adq16646	TPO mimet	237	57	58.8	28	3	ABb17285	Bb17285	TPO-mimet
165	57	58.8	18	8	ADQ16607	Adq16607	TPO mimet	238	57	58.8	28	5	ABP51682	Abp51682	TPO mimet
166	57	58.8	18	8	ADQ16615	Adq16615	TPO mimet	239	57	58.8	28	7	ADJ73013	Adj73013	TPO mimet
167	57	58.8	18	8	ADQ16627	Adq16627	TPO mimet	240	57	58.8	28	8	ADJ52648	Adj52648	CH1 delet
168	57	58.8	18	8	ADQ16617	Adq16617	TPO mimet	241	57	58.8	28	8	ADJ51609	Adj51609	CH1 delet
169	57	58.8	18	8	ADQ16613	Adq16613	TPO mimet	242	57	58.8	28	8	ADQ16636	Adq16636	Tetanus t
170	57	58.8	18	8	ADQ16623	Adq16623	TPO mimet	243	57	58.8	28	9	ADV44374	Adv44374	Modified

244	57	58.8	28	9	AEI12847	Aebi2847 Antibody	317	57	58.8	269	3	AA996531	AA996531 Human IGG
245	57	58.8	29	3	AAI16975	Aabi16975 TPO-mimet	318	57	58.8	269	3	AAI16960	Aabi16960 TPO-mimet
246	57	58.8	29	3	AAI16975	Aabi16975 TPO-mimet	319	57	58.8	269	3	AAI16960	Aabi16960 TPO-mimet
247	57	58.8	29	3	AAI17286	Aabi17286 TPO-mimet	320	57	58.8	472	5	ABP51695	ABP51695 SGI.1-TPO
248	57	58.8	29	3	AAI16970	Aabi16970 TPO-mimet	321	57	58.8	472	5	ADQ16647	ADQ16647 Immunoglob
249	57	58.8	29	3	AAI16970	Aabi16970 TPO-mimet	322	57	58.8	472	9	ADV44385	ADV44385 SGI.1 hea
250	57	58.8	29	5	ABP72862	Abb72862 TPO-mimet	323	57	58.8	472	9	AEV12858	Aebi2858 Antibody
251	57	58.8	29	5	ABP72862	Abb72862 TPO-mimet	324	56	57.7	14	2	AAW36773	AAW36773 Thrombopo
252	57	58.8	29	7	ADJ73011	Adj73011 TPO-mimet	325	56	57.7	14	2	AAW36773	AAW36773 Thrombopo
253	57	58.8	29	7	ADJ73006	Adj73006 TPO-mimet	326	56	57.7	15	8	ADM72527	ADM72527 TPO mimet
254	57	58.8	29	8	ADJ52646	Adj52646 CHI delet	327	56	57.7	15	2	AAW66731	AAW66731 Peptide c
255	57	58.8	29	8	ADJ52641	Adj52641 CHI delet	328	56	57.7	15	4	AAU26038	AAU26038 Human thr
256	57	58.8	29	8	ADJ51602	Adj51602 CHI delet	329	56	57.7	18	7	ADN59663	ADN59663 Thrombopo
257	57	58.8	29	8	ADJ51607	Adj51607 CHI delet	330	56	57.7	22	7	ADN59830	ADN59830 TPO pepti
258	57	58.8	29	8	ADJ51607	Adj51607 TPO-mimet	331	56	57.7	25	7	ADN59708	ADN59708 Thrombopo
259	57	58.8	30	3	AAI17287	Aabi17287 TPO-mimet	332	56	57.7	43	7	ADN59759	ADN59759 Peptide-v
260	57	58.8	31	3	AAI17286	Aabi17286 TPO-mimet	333	54	55.7	13	4	AAU26008	AAU26008 Human thr
261	57	58.8	31	3	AAI16974	Aabi16974 TPO-mimet	334	54	55.7	13	4	AAU26012	AAU26012 Human thr
262	57	58.8	31	3	AAI16973	Aabi16973 TPO-mimet	335	54	55.7	14	3	AAI16969	Aabi16969 TPO-mimet
263	57	58.8	31	5	ABP72860	Abb72860 TPO-mimet	336	54	55.7	14	3	AAI16968	Aabi16968 TPO-mimet
264	57	58.8	31	5	ABP72859	Abb72859 TPO-mimet	337	54	55.7	14	5	ABP72854	Abb72854 TPO mimet
265	57	58.8	31	7	ADJ73009	Adj73009 TPO-mimet	338	54	55.7	14	5	ABP72855	Abb72855 TPO mimet
266	57	58.8	31	7	ADJ73010	Adj73010 TPO-mimet	339	54	55.7	14	7	ADJ73004	Adj73004 CHI delet
267	57	58.8	31	8	ADJ52644	Adj52644 CHI delet	340	54	55.7	14	8	ADJ52639	Adj52639 CHI delet
268	57	58.8	31	8	ADJ52645	Adj52645 CHI delet	341	54	55.7	14	8	ADJ52640	Adj52640 CHI delet
269	57	58.8	31	8	ADJ51606	Adj51606 CHI delet	342	54	55.7	14	8	ADJ51601	Adj51601 CHI delet
270	57	58.8	32	3	AAI169520	Aabi169520 Thrombopo	343	54	55.7	14	8	ADJ51600	Adj51600 CHI delet
271	57	58.8	32	3	AAI17289	Aabi17289 TPO-mimet	344	54	55.7	15	2	AAW66718	AAW66718 Peptide c
272	57	58.8	32	3	AAI17290	Aabi17290 TPO-mimet	345	54	55.7	15	2	AAW66718	AAW66718 Peptide c
273	57	58.8	33	3	AAI17290	Aabi17290 TPO-mimet	346	54	55.7	29	5	ABP72857	Abb72857 TPO mimet
274	57	58.8	34	3	AAI17291	Aabi17291 TPO-mimet	347	54	55.7	29	7	ADJ73007	Adj73007 TPO mimet
275	57	58.8	34	3	AAI17292	Aabi17292 TPO-mimet	348	54	55.7	29	8	ADJ52642	Adj52642 CHI delet
276	57	58.8	35	3	AAI17292	Aabi17292 TPO-mimet	349	54	55.7	29	8	ADJ51603	Adj51603 CHI delet
277	57	58.8	36	3	AAI17292	Aabi17292 TPO-mimet	350	53	54.6	12	2	AAW36787	AAW36787 Thrombopo
278	57	58.8	36	3	AAI17292	Aabi17292 TPO-mimet	351	53	54.6	12	2	AAW36787	AAW36787 Thrombopo
279	57	58.8	36	3	AAI17292	Aabi17292 TPO-mimet	352	53	54.6	13	4	AAU26035	AAU26035 Human thr
280	57	58.8	36	3	AAI17292	Aabi17292 TPO-mimet	353	53	54.6	13	4	AAU26035	AAU26035 Human thr
281	57	58.8	36	3	AAI17307	Aabi17307 TPO-mimet	354	53	54.6	13	7	ADJ73003	Adj73003 TPO mimet
282	57	58.8	36	3	AAI17303	Aabi17303 TPO-mimet	355	53	54.6	13	8	ADJ52638	Adj52638 CHI delet
283	57	58.8	36	3	AAI17303	Aabi17303 TPO-mimet	356	53	54.6	13	8	ADJ51599	Adj51599 CHI delet
284	57	58.8	36	3	AAI16963	Aabi16963 TPO-mimet	357	53	54.6	13	8	ADJ52524	ADJ52524 TPO mimet
285	57	58.8	36	3	AAI17301	Aabi17301 TPO-mimet	358	53	54.6	14	2	AAW36788	AAW36788 Thrombopo
286	57	58.8	36	3	AAI17306	Aabi17306 TPO-mimet	359	53	54.6	14	2	AAW36788	AAW36788 Thrombopo
287	57	58.8	36	5	ABP72403	Abb72403 TPO-mimet	360	53	54.6	14	4	AAU26013	AAU26013 Human thr
288	57	58.8	37	3	AAI17294	Aabi17294 TPO-mimet	361	53	54.6	15	4	AAU26011	AAU26011 Human thr
289	57	58.8	38	3	AAI17295	Aabi17295 TPO-mimet	362	53	54.6	15	4	AAU26011	AAU26011 Human thr
290	57	58.8	39	3	AAI17304	Aabi17304 TPO-mimet	363	53	54.6	19	2	AAW09457	AAW09457 Thrombopo
291	57	58.8	39	3	AAI17305	Aabi17305 TPO-mimet	364	53	54.6	19	2	AAW09457	AAW09457 Thrombopo
292	57	58.8	40	3	AAI17302	Aabi17302 TPO-mimet	365	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
293	57	58.8	41	3	AAI17302	Aabi17302 TPO-mimet	366	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
294	57	58.8	41	5	ABP73388	Abb73388 TPO-mimet	367	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
295	57	58.8	41	5	ABP73388	Abb73388 TPO-mimet	368	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
296	57	58.8	42	3	AAI17296	Aabi17296 TPO-mimet	369	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
297	57	58.8	42	3	AAI17296	Aabi17296 TPO-mimet	370	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
298	57	58.8	42	3	AAI17282	Aabi17282 TPO-mimet	371	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
299	57	58.8	42	3	AAI17281	Aabi17281 TPO-mimet	372	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
300	57	58.8	42	3	AAI173404	Abb73404 TPO-mimet	373	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
301	57	58.8	42	5	ABP73388	Abb73388 TPO-mimet	374	53	54.6	19	2	AAW36651	AAW36651 Thrombopo
302	57	58.8	60	3	AAI17311	Aabi17311 Synthetic	375	52	53.6	12	2	AAW36781	AAW36781 Thrombopo
303	57	58.8	60	3	AAI17311	Aabi17311 Synthetic	376	52	53.6	12	2	AAW36781	AAW36781 Thrombopo
304	57	58.8	122	9	ADV44474	Adv44474 Tmp-TMP g	377	52	53.6	13	4	AAU26018	AAU26018 Human thr
305	57	58.8	122	9	ADV44474	Adv44474 Tmp-TMP g	378	52	53.6	13	4	AAU26018	AAU26018 Human thr
306	57	58.8	129	6	ABG71751	Abg71751 Antibody	379	52	53.6	13	8	ADM72525	ADM72525 TPO mimet
307	57	58.8	131	6	ABG71753	Abg71753 Antibody	380	52	53.6	13	8	ADM72488	ADM72488 TPO mimet
308	57	58.8	133	6	ABG71752	Abg71752 Antibody	381	52	53.6	13	8	ADM72488	ADM72488 TPO mimet
309	57	58.8	135	6	ABG71749	Abg71749 Antibody	382	52	53.6	13	8	ADM72489	ADM72489 TPO mimet
310	57	58.8	143	6	ABG71750	Abg71750 Antibody	383	52	53.6	14	2	AAW66715	AAW66715 Peptide c
311	57	58.8	247	3	AAI16958	Aabi16958 Fc-TMP pr	384	52	53.6	14	2	AAW66715	AAW66715 Peptide c
312	57	58.8	247	3	AAI16958	Aabi16958 Fc-TMP pr	385	52	53.6	14	2	AAW66715	AAW66715 Peptide c
313	57	58.8	247	5	ABP73411	Abb73411 Fc-TPO mi	386	52	53.6	14	2	AAW66715	AAW66715 Peptide c
314	57	58.8	247	5	ABP73411	Abb73411 Fc-TPO mi	387	52	53.6	14	2	AAW66715	AAW66715 Peptide c
315	57	58.8	268	3	AAI16959	Aabi16959 Fc-TMP-TM	388	52	53.6	15	4	AAU26022	AAU26022 Human thr
316	57	58.8	268	5	ABP73412	Abb73412 Fc-TMP-TM	389	52	53.6	15	4	AAU26022	AAU26022 Human thr

390	52	53.6	15	8	ADM72496	Adm72496	TPO mimet	463	47	48.5	8	5	ABP51678	Abp51678	TPO mimet
391	52	53.6	15	8	ADM72494	Adm72494	TPO mimet	464	47	48.5	8	8	ADQ16692	Adq16692	TPO mimet
392	52	53.6	18	7	ADNS9680	Adns9681	Thrombopo	465	47	48.5	8	9	ADV44430	Adv44430	Agonist t
393	52	53.6	18	7	ADNS9667	Adns9667	Thrombopo	466	47	48.5	8	9	AEAB12902	Aeb12902	TPO mimet
394	52	53.6	19	2	AAW09494	Aw09494	Thrombopo	467	47	48.5	14	8	ADM72509	Adm72509	TPO mimet
395	52	53.6	18	2	AAW09461	Aw09461	Thrombopo	468	47	48.5	15	2	AAW66723	Aaw66723	Peptide c
396	52	53.6	19	2	AAW09461	Aw09461	Thrombopo	469	47	48.5	15	4	Aau26028	Aau26028	Human thr
397	52	53.6	19	2	AAW33028	Aw33028	Thrombopo	470	47	48.5	15	4	ADM72508	Adm72508	Human thr
398	52	53.6	19	2	AAW36645	Aw36645	Thrombopo	471	47	48.5	15	4	ABP51679	Abp51679	TPO mimet
399	52	53.6	19	3	AAAB17022	Aab17022	TPO-mimet	472	47	48.5	18	5	ADNS59672	Adns59672	Thrombopo
400	52	53.6	19	4	AAU25864	Aau25864	Human thr	473	47	48.5	18	7	ADNS59658	Adns59658	Thrombopo
401	52	53.6	19	4	AAU25825	Aau25825	Human thr	474	47	48.5	18	7	ADNS59675	Adns59675	Thrombopo
402	52	53.6	19	5	ABB72908	Abb72908	TPO mimet	475	47	48.5	18	7	ADNS59678	Adns59678	Thrombopo
403	52	53.6	19	8	ADJ73060	Adj73060	TPO mimet	476	47	48.5	18	7	ADNS59661	Adns59661	Thrombopo
404	52	53.6	19	8	ADJ52695	Adj52695	CHI delet	477	47	48.5	18	8	ADQ16693	Adq16693	TPO mimet
405	52	53.6	19	8	ADJ51656	Adj51656	CHI delet	478	47	48.5	18	9	ADV44431	Adv44431	Anti-teta
406	52	53.6	22	7	ADNS59834	Adns59834	TMP pepti	479	47	48.5	18	9	AEAB12903	Aeb12903	Heavy cha
407	52	53.6	25	7	ADNS59742	Adns59742	Thrombopo	480	47	48.5	22	7	ADNS59828	Adns59828	TMP pepti
408	52	53.6	25	7	ADNS59716	Adns59716	Thrombopo	481	47	48.5	22	7	ADNS59825	Adns59825	TMP pepti
409	52	53.6	25	7	ADNS59740	Adns59740	Thrombopo	482	47	48.5	22	7	ADNS59839	Adns59839	TMP pepti
410	52	53.6	42	7	ADNS9818	Adns9818	Peptide-	483	47	48.5	25	7	ADNS59736	Adns59736	Thrombopo
411	52	53.6	43	7	ADNS9752	Adns9752	Peptide-v	484	47	48.5	25	7	ADNS59704	Adns59704	Thrombopo
412	51	52.6	15	2	AAW66719	Aaw66719	Peptide c	485	47	48.5	25	7	ADNS59730	Adns59730	Thrombopo
413	50	51.5	12	4	AAU26014	Aau26014	Human thr	486	47	48.5	25	7	ADNS59698	Adns59698	Thrombopo
414	49	50.5	18	7	ADNS59815	Adns59815	Thrombopo	487	47	48.5	25	7	ADNS59744	Adns59744	Thrombopo
415	49	50.5	18	7	ADNS9653	Adns9653	Thrombopo	488	47	48.5	40	7	ADNS59753	Adns59753	Peptide-v
416	49	50.5	18	7	ADNS9669	Adns9669	Thrombopo	489	46.5	47.9	341	6	AAU33868	Aau33868	Protein e
417	49	50.5	18	7	ADNS9673	Adns9673	Thrombopo	490	46	47.4	10	2	AAW09469	Aaw09469	Thrombopo
418	49	50.5	22	7	ADNS9820	Adns9820	TMP pepti	491	46	47.4	10	2	AAW36621	Aaw36621	Thrombopo
419	49	50.5	22	7	ADNS9840	Adns9840	TMP pepti	492	46	47.4	10	2	AAW36772	Aaw36772	Thrombopo
420	49	50.5	22	7	ADNS9836	Adns9836	TMP pepti	493	46	47.4	10	3	AAAB17005	Aab17005	TPO-mimet
421	49	50.5	23	7	ADNS9778	Adns9778	Peptide-v	494	46	47.4	10	4	AAU25839	Aau25839	Human thr
422	49	50.5	23	7	ADNS9796	Adns9796	Peptide-v	495	46	47.4	10	4	AAU25840	Aau25840	Human thr
423	49	50.5	25	7	ADNS9726	Adns9726	Thrombopo	496	46	47.4	10	5	ABB72891	Abb72891	TPO mimet
424	49	50.5	25	7	ADNS9689	Adns9689	Thrombopo	497	46	47.4	10	7	ADJ73042	Adj73042	TPO mimet
425	49	50.5	25	7	ADNS9720	Adns9720	Thrombopo	498	46	47.4	10	8	ADJ52677	Adj52677	CHI delet
426	49	50.5	36	7	ADNS9766	Adns9766	Peptide-v	499	46	47.4	10	8	ADJ51638	Adj51638	CHI delet
427	49	50.5	36	7	ADNS9762	Adns9762	Peptide-v	500	46	47.4	14	8	ADM72499	Adm72499	TPO mimet
428	49	50.5	36	7	ADNS9756	Adns9756	Peptide-v	501	46	47.4	15	4	AAU26024	Aau26024	Human thr
429	49	50.5	41	7	ADNS9816	Adns9816	Peptide-	502	46	47.4	15	8	ADM72498	Adm72498	TPO mimet
430	49	50.5	41	7	ADNS9772	Adns9772	Peptide-v	503	46	47.4	18	2	AAW36638	Aaw36638	Thrombopo
431	49	50.5	46	7	ADNS9790	Adns9790	Peptide-v	504	46	47.4	18	2	AAW09487	Aaw09487	Thrombopo
432	49	50.5	46	7	ADNS9784	Adns9784	Peptide-v	505	46	47.4	18	4	Aau25857	Aau25857	Human thr
433	48	49.5	10	3	AAAB17006	Aab17006	TPO-mimet	506	46	47.4	18	7	ADNS59866	Adns59866	Thrombopo
434	48	49.5	10	5	ABB72892	Abb72892	TPO mimet	507	46	47.4	22	7	ADNS59833	Adns59833	TMP pepti
435	48	49.5	10	8	ADJ73043	Adj73043	TPO mimet	508	46	47.4	25	7	ADNS59714	Adns59714	Thrombopo
436	48	49.5	10	8	ADJ52678	Adj52678	CHI delet	509	45	46.4	13	3	AAAB17015	Aab17015	TPO-mimet
437	48	49.5	10	8	ADJ51639	Adj51639	CHI delet	510	45	46.4	13	5	ABB72901	Abb72901	TPO mimet
438	48	49.5	12	3	AAAB17309	Aab17309	Synthetic	511	45	46.4	13	5	ADJ73054	Adj73054	TPO mimet
439	48	49.5	13	2	AAW36792	Aaw36792	Thrombopo	512	45	46.4	13	7	ADJ73052	Adj73052	TPO mimet
440	48	49.5	13	4	AAU26015	Aau26015	Human thr	513	45	46.4	13	7	ADJ73056	Adj73056	TPO mimet
441	48	49.5	18	7	ADNS9660	Adns9660	Thrombopo	514	45	46.4	13	7	ADJ73053	Adj73053	TPO mimet
442	48	49.5	18	7	ADNS9654	Adns9654	Thrombopo	515	45	46.4	13	7	ADJ73055	Adj73055	TPO mimet
443	48	49.5	18	7	ADNS9655	Adns9655	Thrombopo	516	45	46.4	13	8	ADJ52687	Adj52687	CHI delet
444	48	49.5	18	7	ADNS9670	Adns9670	Thrombopo	517	45	46.4	13	8	ADJ51648	Adj51648	CHI delet
445	48	49.5	22	7	ADNS9821	Adns9821	TMP pepti	518	45	46.4	14	3	AAAB17017	Aab17017	TPO-mimet
446	48	49.5	22	7	ADNS9827	Adns9827	TMP pepti	519	45	46.4	14	3	AAAB17016	Aab17016	TPO-mimet
447	48	49.5	22	7	ADNS9822	Adns9822	TMP pepti	520	45	46.4	14	5	ABB72903	Abb72903	TPO mimet
448	48	49.5	22	7	ADNS9837	Adns9837	TMP pepti	521	45	46.4	14	5	ABB72902	Abb72902	TPO mimet
449	48	49.5	23	7	ADNS9792	Adns9792	Peptide-v	522	45	46.4	14	8	ADJ52689	Adj52689	CHI delet
450	48	49.5	23	7	ADNS9774	Adns9774	Peptide-v	523	45	46.4	14	8	ADJ52688	Adj52688	CHI delet
451	48	49.5	25	7	ADNS9722	Adns9722	Thrombopo	524	45	46.4	14	8	ADJ51650	Adj51650	CHI delet
452	48	49.5	25	7	ADNS9692	Adns9692	Thrombopo	525	45	46.4	14	8	ADJ51649	Adj51649	CHI delet
453	48	49.5	25	7	ADNS9702	Adns9702	Thrombopo	526	45	46.4	15	3	AAAB17018	Aab17018	TPO-mimet
454	48	49.5	25	7	ADNS9691	Adns9691	Thrombopo	527	45	46.4	15	5	ABB72904	Abb72904	TPO mimet
455	48	49.5	25	7	ADNS9768	Adns9768	Peptide-v	528	45	46.4	15	8	ADJ52691	Adj52691	CHI delet
456	48	49.5	42	7	ADNS9751	Adns9751	Peptide-v	529	45	46.4	15	8	ADJ52690	Adj52690	CHI delet
457	48	49.5	43	7	ADNS9761	Adns9761	Peptide-v	530	45	46.4	15	8	ADJ51652	Adj51652	CHI delet
458	48	49.5	44	7	ADNS9817	Adns9817	Peptide-	531	45	46.4	15	8	ADJ51651	Adj51651	CHI delet
459	48	49.5	46	7	ADNS9780	Adns9780	Peptide-v	532	45	46.4	18	7	ADNS59664	Adns59664	Thrombopo
460	48	49.5	46	7	ADNS9786	Adns9786	Peptide-v	533	45	46.4	18	7	ADNS59665	Adns59665	Thrombopo
461	47	48.5	8	2	AAW33037	Aaw33037	Thrombopo	534	45	46.4	18	7	ADNS59671	Adns59671	Thrombopo
462	47	48.5	8	4	AAU25982	Aau25982	Human thr	535	45	46.4	18	7	ADNS59668	Adns59668	Thrombopo

536	45	46.4	18	7	ADN59674	Thrombopo	609	44	45.4	30	9	ADY64336	Thrombopo
537	45	46.4	18	7	ADN59659	Thrombopo	610	44	45.4	36	3	AAV96521	Cyclic or
538	45	46.4	18	7	ADN59657	Thrombopo	611	44	45.4	36	3	AAV17298	TPO-mimet
539	45	46.4	22	7	ADN59826	TMP pepti	612	44	45.4	36	3	AAV17299	TPO-mimet
540	45	46.4	22	7	ADN59824	TMP pepti	613	44	45.4	52	4	AAU56008	Propionib
541	45	46.4	22	7	ADN59841	TMP pepti	614	44	45.4	52	6	ABM52527	Propionib
542	45	46.4	22	7	ADN59835	TMP pepti	615	44	45.4	100	5	ABM52527	Propionib
543	45	46.4	22	7	ADN59831	TMP pepti	616	44	45.4	347	5	ABP65398	Blifidobac
544	45	46.4	22	7	ADN59838	TMP pepti	617	44	45.4	429	5	ABP47758	Human rep
545	45	46.4	22	7	ADN59832	TMP pepti	618	43	45.4	429	5	AAU94594	Human rep
546	45	46.4	22	7	ADN59776	Peptide-v	619	43	44.3	13	4	AAU26039	Human thr
547	45	46.4	23	7	ADN59797	Peptide-v	620	43	44.3	13	4	ADN72528	TPO mimet
548	45	46.4	23	7	ADN59775	Peptide-v	621	43	44.3	18	2	AAU9488	Thrombopo
549	45	46.4	23	7	ADN59779	Peptide-v	622	43	44.3	18	2	AAU36639	Thrombopo
550	45	46.4	23	7	ADN59794	Peptide-v	623	43	44.3	18	4	AAU25858	Human thr
551	45	46.4	23	7	ADN59795	Peptide-v	624	43	44.3	36	3	AAV96522	Linear th
552	45	46.4	23	7	ADN59793	Peptide-v	625	43	44.3	36	3	AAV17300	TPO-mimet
553	45	46.4	23	7	ADN59777	Peptide-v	626	43	44.3	178	4	AAU40877	Propionib
554	45	46.4	25	7	ADN59696	Thrombopo	627	43	44.3	178	6	ABM37396	Propionib
555	45	46.4	25	7	ADN59710	Thrombopo	628	43	44.3	269	6	ABU36398	Protein e
556	45	46.4	25	7	ADN59724	Thrombopo	629	43	44.3	298	3	AGI16591	Arabidops
557	45	46.4	25	7	ADN59712	Thrombopo	630	43	44.3	316	3	AGI16590	Arabidops
558	45	46.4	25	7	ADN59718	Thrombopo	631	43	44.3	369	6	ABU28974	Protein e
559	45	46.4	25	7	ADN59728	Thrombopo	632	43	44.3	375	4	AAU35335	Enterococ
560	45	46.4	25	7	ADN59755	Peptide-v	633	43	44.3	536	3	AGI16589	Arabidops
561	45	46.4	25	7	ADN59700	Thrombopo	634	43	44.3	536	8	ABN93535	Herbicida
562	45	46.4	35	7	ADN59754	Peptide-v	635	43	44.3	536	8	ADN74161	Thale cre
563	45	46.4	36	7	ADN59763	Peptide-v	636	43	44.3	1150	4	AAE11018	Recombina
564	45	46.4	36	7	ADN59765	Peptide-v	637	43	44.3	1150	5	ABB08110	FIV-Oma3
565	45	46.4	36	7	ADN59764	Peptide-v	638	42	43.3	11	2	AAW35425	Thrombopo
566	45	46.4	36	7	ADN59767	Peptide-v	639	42	43.3	11	4	AAU26001	Human thr
567	45	46.4	41	7	ADN59769	Peptide-v	640	42	43.3	12	2	AAW35423	Thrombopo
568	45	46.4	41	7	ADN59770	Peptide-v	641	42	43.3	12	4	AAU26000	Human thr
569	45	46.4	41	7	ADN59771	Peptide-v	642	42	43.3	12	8	ADM72530	TPO mimet
570	45	46.4	41	7	ADN59773	Peptide-v	643	42	43.3	13	2	AAW09467	Thrombopo
571	45	46.4	46	7	ADN59788	Peptide-v	644	42	43.3	13	2	AAW35399	Thrombopo
572	45	46.4	46	7	ADN59789	Peptide-v	645	42	43.3	13	2	AAW35404	Thrombopo
573	45	46.4	46	7	ADN59781	Peptide-v	646	42	43.3	13	2	AAW35417	Thrombopo
574	45	46.4	46	7	ADN59787	Peptide-v	647	42	43.3	13	2	AAW33033	Thrombopo
575	45	46.4	46	7	ADN59757	Peptide-v	648	42	43.3	13	2	AAW35413	Thrombopo
576	45	46.4	46	7	ADN59785	Peptide-v	649	42	43.3	13	2	AAW35406	Thrombopo
577	45	46.4	46	7	ADN59782	Peptide-v	650	42	43.3	13	2	AAW35405	Thrombopo
578	45	46.4	46	7	ADN59783	Peptide-v	651	42	43.3	13	2	AAW35422	Thrombopo
579	45	46.4	46	7	ADN59791	Peptide-v	652	42	43.3	13	2	AAW35397	Thrombopo
580	45	46.4	75	7	ADN59758	Peptide-v	653	42	43.3	13	2	AAW36783	Thrombopo
581	45	46.4	302	6	ABU33828	Protein e	654	42	43.3	13	4	AAU26041	Human thr
582	44	45.4	14	8	ADM72507	TPO mimet	655	42	43.3	13	4	AAU25997	Human thr
583	44	45.4	14	8	ADM72505	TPO mimet	656	42	43.3	13	4	AAU25997	Human thr
584	44	45.4	14	8	ADM72475	TPO mimet	657	42	43.3	13	4	AAU25990	Human thr
585	44	45.4	14	8	ADM72501	TPO mimet	658	42	43.3	13	4	AAU25984	Human thr
586	44	45.4	14	8	ADM72502	Thrombopo	659	42	43.3	13	8	ADM72529	TPO mimet
587	44	45.4	14	9	ADU70205	Thrombopo	660	42	43.3	14	2	AAW09466	Thrombopo
588	44	45.4	14	9	ADY64334	Thrombopo	661	42	43.3	14	2	AAW09462	Thrombopo
589	44	45.4	15	2	AAW66722	Peptide c	662	42	43.3	14	2	AAW09465	Thrombopo
590	44	45.4	15	2	AAW66720	Peptide c	663	42	43.3	14	2	AAW09482	Thrombopo
591	44	45.4	15	4	AAU25833	Human thr	664	42	43.3	14	2	AAW33031	Thrombopo
592	44	45.4	15	4	AAU25833	Human thr	665	42	43.3	14	2	AAW33031	Thrombopo
593	44	45.4	15	4	AAU26027	Human thr	666	42	43.3	14	2	AAW35407	Thrombopo
594	44	45.4	15	4	AAU26025	Human thr	667	42	43.3	14	2	AAW35407	Thrombopo
595	44	45.4	15	8	ADM72506	TPO mimet	668	42	43.3	14	2	AAW35408	Thrombopo
596	44	45.4	15	8	ADM72500	TPO mimet	669	42	43.3	14	2	AAW35398	Thrombopo
597	44	45.4	15	8	ADM72477	TPO mimet	670	42	43.3	14	2	AAW36633	Thrombopo
598	44	45.4	15	8	ADM72504	TPO mimet	671	42	43.3	14	2	AAW33029	Thrombopo
599	44	45.4	15	9	ADY64335	Thrombopo	672	42	43.3	14	2	AAW35396	Thrombopo
600	44	45.4	16	2	AAW66711	Peptide c	673	42	43.3	14	2	AAW35403	Thrombopo
601	44	45.4	16	8	ADM72476	TPO mimet	674	42	43.3	14	2	AAW35403	Thrombopo
602	44	45.4	18	8	ADJ52647	CH1 delet	675	42	43.3	14	2	AAW36647	Thrombopo
603	44	45.4	29	3	ABM16972	TPO-mimet	676	42	43.3	14	2	AAW35400	Thrombopo
604	44	45.4	29	5	ABM72858	TPO mimet	677	42	43.3	14	2	AAW35402	Thrombopo
605	44	45.4	29	7	ADJ73012	TPO mimet	678	42	43.3	14	2	AAW33032	Thrombopo
606	44	45.4	29	8	ADJ73008	TPO mimet	679	42	43.3	14	2	AAW66732	Peptide c
607	44	45.4	29	8	ADJ52643	CH1 delet	680	42	43.3	14	3	AAW17010	TPO-mimet
608	44	45.4	29	8	ADJ51604	CH1 delet	681	42	43.3	14	3	AAW17014	TPO-mimet
					ADJ51608	CH1 delet							

682	42	43.3	14	4	AAU25987	Aau25987 Human thr	755	41	42.3	18	4	AAU25867	Aau25867 Human thr
683	42	43.3	14	4	AAU25826	Aau25826 Human thr	756	41	42.3	18	5	ABB72911	Abb72911 TPO mimet
684	42	43.3	14	4	AAU25993	Aau25993 Human thr	757	41	42.3	18	7	ADJ73063	Adj73063 TPO mimet
685	42	43.3	14	4	AAU25852	Aau25852 Human thr	758	41	42.3	18	7	ADN59679	Adn59679 Thrombopo
686	42	43.3	14	4	AAU25989	Aau25989 Human thr	759	41	42.3	18	7	ADN59662	Adn59662 Thrombopo
687	42	43.3	14	4	AAU25983	Aau25983 Human thr	760	41	42.3	18	8	ADJ52698	Adj52698 CHI delet
688	42	43.3	14	4	AAU25985	Aau25985 Human thr	761	41	42.3	18	8	ADJ51659	Adj51659 CHI delet
689	42	43.3	14	4	AAU25995	Aau25995 Human thr	762	41	42.3	18	8	ADN59829	Adn59829 TWP pept1
690	42	43.3	14	4	AAU25992	Aau25992 Human thr	763	41	42.3	25	7	ADN59738	Adn59738 Thrombopo
691	42	43.3	14	4	AAU25986	Aau25986 Human thr	764	41	42.3	25	7	ADN59706	Adn59706 Thrombopo
692	42	43.3	14	4	AAU26040	Aau26040 Human thr	765	41	42.3	40	7	ADN59760	Adn59760 Peptide-v
693	42	43.3	14	4	AAU25866	Aau25866 Human thr	766	41	42.3	70	7	ADC88216	Adc88216 Ribosomal
694	42	43.3	14	4	AAU25988	Aau25988 Human thr	767	41	42.3	99	5	ABP47803	Abp47803 FIV prote
695	42	43.3	14	5	ABB72896	Abb72896 TPO mimet	768	41	42.3	184	7	ADH88208	Adh88208 Enterococ
696	42	43.3	14	5	ABBY2900	Abby2900 TPO mimet	769	41	42.3	187	3	AAB39464	Aab39464 Gene 26 h
697	42	43.3	14	7	ADJ73047	Adj73047 TPO mimet	770	41	42.3	229	3	AAG34219	Aag34219 Zea may
698	42	43.3	14	7	ADJ73051	Adj73051 TPO mimet	771	41	42.3	244	5	ABG66710	Abg66710 Human nov
699	42	43.3	14	8	ADJ52682	Adj52682 CHI delet	772	41	42.3	263	5	ABP40153	Abp40153 Staphyloc
700	42	43.3	14	8	ADJ52686	Adj52686 CHI delet	773	41	42.3	263	8	ADN507199	Adn507199 Staphyloc
701	42	43.3	14	8	ADJ51643	Adj51643 CHI delet	774	41	42.3	266	7	ADS92500	Ads92500 B. lichen
702	42	43.3	14	8	ADJ51647	Adj51647 CHI delet	775	41	42.3	276	6	ABU31149	Abu31149 Protein e
703	42	43.3	18	2	AAW09456	Aaw09456 Thrombopo	776	41	42.3	283	8	ADA2544	Ada2544 Bacterial
704	42	43.3	18	2	AAW33023	Aaw33023 Thrombopo	777	41	42.3	295	3	AAG34218	Aag34218 Zea may
705	42	43.3	18	3	AAB17020	Aab17020 TPO-mimet	778	41	42.3	306	5	ABG91624	Abg91624 Purine/py
706	42	43.3	18	4	AAU25820	Aau25820 Human thr	779	41	42.3	306	6	ABU36556	Abu36556 Protein e
707	42	43.3	18	5	ABBY2905	Abby2905 TPO mimet	780	41	42.3	306	6	ADT57646	Adt57646 Plant pol
708	42	43.3	18	7	ADJ73058	Adj73058 TPO mimet	781	41	42.3	329	8	ADT57646	Adt57646 Plant pol
709	42	43.3	18	7	ADN59677	Adn59677 Thrombopo	782	41	42.3	370	8	ADY11497	Ady11497 Plant ful
710	42	43.3	18	7	ADN59676	Adn59676 Thrombopo	783	41	42.3	376	8	ABM84818	Abm84818 Human dia
711	42	43.3	18	7	ADN59656	Adn59656 Thrombopo	784	41	42.3	376	8	ABM84819	Abm84819 Human dia
712	42	43.3	18	8	ADJ52693	Adj52693 CHI delet	785	41	42.3	406	3	AAB54065	Aab54065 Human pan
713	42	43.3	18	8	ADJ51654	Adj51654 CHI delet	786	41	42.3	419	2	AAW01508	Aaw01508 Human pan
714	42	43.3	19	2	AAW09458	Aaw09458 Thrombopo	787	41	42.3	419	2	AAW01509	Aaw01509 Human pan
715	42	43.3	19	2	AAW33025	Aaw33025 Thrombopo	788	41	42.3	419	2	AAW01504	Aaw01504 Wild-type
716	42	43.3	19	4	AAU25822	Aau25822 Human thr	789	41	42.3	419	2	AAW01510	Aaw01510 Human pan
717	42	43.3	22	7	ADN59823	Adn59823 TWP pept1	790	41	42.3	419	2	AAU25815	Aau25815 Human car
718	42	43.3	25	4	AAU26042	Aau26042 Human thr	791	41	42.3	419	2	AAU258915	Aau258915 Human reg
719	42	43.3	25	7	ADN59694	Adn59694 Thrombopo	792	41	42.3	419	5	ADI17250	Adi17250 Human NOV
720	42	43.3	25	7	ADN59734	Adn59734 Thrombopo	793	41	42.3	419	5	ADI17249	Adi17249 Human NOV
721	42	43.3	25	7	ADN59732	Adn59732 Thrombopo	794	41	42.3	419	6	ABR54273	Abr54273 Human NOV
722	42	43.3	25	8	ADM72531	Adm72531 TPO mimet	795	41	42.3	419	6	ABR54276	Abr54276 Human NOV
723	42	43.3	57	4	ABG03094	Abg03094 Novel hum	796	41	42.3	419	6	ABR54276	Abr54276 Human car
724	42	43.3	57	4	ABG03096	Abg03096 Novel hum	797	41	42.3	419	8	ADQ30585	Adq30585 Pancreas
725	42	43.3	81	7	ABO75974	AbO75974 Pseudomon	798	41	42.3	426	6	ABR54275	Abr54275 Human NOV
726	42	43.3	140	3	ABR18086	Aab18086 Pinus rad	799	41	42.3	431	5	ABP47760	Abp47760 Protein #
727	42	43.3	180	8	ADX90944	Adx90944 Plant ful	800	41	42.3	432	7	ABO61826	AbO61826 Klebsiell
728	42	43.3	221	7	ABO83112	AbO83112 Pseudomon	801	41	42.3	545	8	ADY12442	Ady12442 Plant ful
729	42	43.3	261	7	ADH86920	Adh86920 Enterococ	802	41	42.3	548	8	ADY08486	Ady08486 Plant ful
730	42	43.3	318	9	ABM91698	Abm91698 M. xanthu	803	41	42.3	559	6	ABU40498	Abu40498 Bacterial
731	42	43.3	320	8	ADX78492	Adx78492 Plant ful	804	41	42.3	572	7	ADF06852	Adf06852 Chicken g
732	42	43.3	328	8	ADX68190	Adx68190 Plant ful	805	41	42.3	573	7	ADP95091	Adp95091 Chicken g
733	42	43.3	470	9	ABM92990	Abm92990 M. xanthu	806	41	42.3	573	8	ADU59659	Adu59659 Guanilate
734	42	43.3	555	7	ADF05437	Adf05437 Bacterial	807	41	42.3	721	5	ABP26887	Abp26887 Streptoco
735	42	43.3	619	7	ADH87640	Adh87640 Enterococ	808	41	42.3	721	6	ABU46751	Abu46751 Protein e
736	42	43.3	806	4	AAW51617	Aam51617 Human PDE	809	41	42.3	721	9	ABR91543	Abr91543 Microbial
737	42	43.3	1061	2	AAW04192	Aar04192 Pol gene	810	41	42.3	762	4	ABG08032	Abg08032 Novel hum
738	42	43.3	1239	6	ABU18757	Abu18757 Protein e	811	41	42.3	1055	3	ABR12993	Abr12993 HIV-2 rev
739	42	43.3	1241	6	ABU18092	Abu18092 Protein e	812	41	42.3	3437	4	ABR66639	Abbr66639 Drosophil
740	42	43.3	4051	7	ADU77465	Adm77465 Rat fibro	813	40.5	41.8	520	7	ABO64867	AbO64867 Klebsiell
741	41.5	42.8	110	4	AAW85655	Aam85655 Human imm	814	40.5	41.8	569	7	ADC68464	Adc68464 Lolium b
742	41.5	42.8	168	4	ABB65448	Abb65448 Drosophil	815	40.5	41.8	569	9	ABO3167	AbO3167 Fructan b
743	41.5	42.8	252	6	ABU17253	Abu17253 Protein e	816	40.5	41.8	571	5	ABG69061	Abg69061 Amino aci
744	41.5	42.8	262	6	ADA34120	Ada34120 Acinetoba	817	40	41.2	15	9	ADV21762	Adv21762 SIV pol p
745	41.5	42.8	308	4	AAG91969	Aag91969 C glucami	818	40	41.2	18	2	AAW09490	Aaw09490 Thrombopo
746	41	42.3	10	2	AAW09472	Aaw09472 Thrombopo	819	40	41.2	18	2	AAW36641	Aaw36641 Thrombopo
747	41	42.3	10	2	AAW36623	Aaw36623 Thrombopo	820	40	41.2	18	4	AAU25860	Aau25860 Human thr
748	41	42.3	10	4	AAU25842	Aau25842 Human thr	821	40	41.2	19	2	AAW09495	Aaw09495 Thrombopo
749	41	42.3	15	2	AAW66724	Aaw66724 Peptide c	822	40	41.2	19	2	AAW36646	Aaw36646 Thrombopo
750	41	42.3	18	2	AAW09497	Aaw09497 Thrombopo	823	40	41.2	19	3	ABR17023	Abr17023 TPO-mimet
751	41	42.3	18	2	AAW36653	Aaw36653 Thrombopo	824	40	41.2	19	4	AAU25865	Aau25865 Human thr
752	41	42.3	18	2	AAW36648	Aaw36648 Thrombopo	825	40	41.2	19	5	ABR72909	Abbr72909 TPO mimet
753	41	42.3	18	3	ABR17025	Aab17025 TPO-mimet	826	40	41.2	19	7	ADJ73061	Adj73061 TPO mimet
754	41	42.3	18	4	AAU25872	Aau25872 Human thr	827	40	41.2	19	8	ADJ52696	Adj52696 CHI delet

828	40	41.2	40	1060	2	AAW89314	AAW89314	STVmac239
829	40	41.2	40	1061	8	ABM82874	ABM82874	Human dia
830	40	41.2	40	1078	4	AAW40012	AAW40012	Human pol
831	40	41.2	40	1104	4	ABW67107	ABW67107	Drosophila
832	40	41.2	40	1118	9	ADV23765	ADV23765	SIV full
833	40	41.2	40	1712	5	AAO17361	AAO17361	Human the
834	40	41.2	40	1712	8	ADU06526	ADU06526	Human the
835	40	41.2	40	1712	8	ADU06526	ADU06526	Human the
836	40	41.2	40	1712	8	ADU06526	ADU06526	Human the
837	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
838	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
839	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
840	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
841	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
842	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
843	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
844	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
845	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
846	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
847	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
848	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
849	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
850	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
851	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
852	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
853	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
854	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
855	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
856	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
857	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
858	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
859	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
860	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
861	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
862	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
863	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
864	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
865	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
866	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
867	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
868	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
869	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
870	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
871	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
872	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
873	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
874	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
875	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
876	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
877	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
878	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
879	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
880	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
881	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
882	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
883	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
884	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
885	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
886	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
887	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
888	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
889	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
890	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
891	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
892	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
893	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
894	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
895	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
896	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
897	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
898	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
899	40	41.2	40	10625	9	ADY59885	ADY59885	Human col
900	40	41.2	40	10625	9	ADY59885	ADY59885	Human col

974 39 40.2 678 3 AAG45616 Arabidops
 975 39 40.2 686 8 ABM84061 Human dia
 976 39 40.2 699 9 ADV85513 Human tra
 977 39 40.2 721 9 ADZ26477 Human ery
 978 39 40.2 747 7 ADB64291 Human pro
 979 39 40.2 806 3 AAG45615 Arabidops
 980 39 40.2 826 3 AAG45614 Arabidops
 981 39 40.2 836 4 ABB68771 Arabidops
 982 39 40.2 839 3 AAG30787 Arabidops
 983 39 40.2 859 3 AAG30786 Arabidops
 984 39 40.2 899 7 AAG73063 Pseudomon
 985 39 40.2 899 7 ADL12118 Pseudomon
 986 39 40.2 905 7 ADM04443 Human pro
 987 39 40.2 944 4 ABB59676 Arabidops
 988 39 40.2 946 4 ABB72016 Arabidops
 989 39 40.2 999 9 ADX40028 HIV Pol p
 990 39 40.2 1057 4 ABG09834 Novel hum
 991 39 40.2 1395 4 ABG04125 Novel hum
 992 39 40.2 1544 5 AAU97541 Human pho
 993 38.5 39.7 193 7 ADM25416 Hyperther
 994 38.5 39.7 258 6 ABM72323 Staphyloc
 995 38.5 39.7 439 5 ADR32246 Human tum
 996 38.5 39.7 439 6 ADA54985 Human pro
 997 38.5 39.7 520 4 AAG92227 C glutami
 998 38.5 39.7 520 7 ADD13669 C. glutam
 999 38.5 39.7 572 7 ABO80566 Pseudomon
 1000 38.5 39.7 2476 2 AAW67738 Pig p105

ALIGNMENTS

RESULT 1
 AAW09499
 ID AAW09499 standard; protein; 18 AA.

XX AC AAW09499;
 XX DT 10-SEP-1997 (first entry)
 XX

XX Thrombopoietin receptor binding peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 XX bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX WO9640189-A1.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PA (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

XX WPI; 1997-051883/05.

XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
 PT mimetic(s) - useful in treatment of haematological disorders, esp.

XX thrombocytopenia resulting from chemotherapy, etc.

XX Disclosure; Page 27; 106pp; English.

CC The present sequence is a peptide which binds to thrombopoietin (TPO)
 CC receptor (TR). The compound can be used for treating patients suffering
 CC from haematological disorders and thrombocytopenia resulting from
 CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
 CC may also be used to maintain the proliferation and growth of TPO-
 CC dependent cell lines and for use in biological research, for detecting
 CC TPO receptors on living cells

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSRHTS 18

DB 1 TIKGPTLRQWLKSRHTS 18

RESULT 2

AAW09459
 ID AAW09459 standard; protein; 18 AA.

XX AC AAW09459;

XX DT 10-SEP-1997 (first entry)

XX Thrombopoietin receptor binding compound peptide.

XX Haematology; thrombocytopenia; TPO; TR; proliferation;
 XX bone marrow transfusion; chemotherapy; radiation therapy.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 1..18

FT /note= "Preferably linkages are selected from: -
 FT CH2OC(O)NR-; phosphonate; -CH2S(O)2NR-; -CH2NR-; -C(O)NR6
 FT ; -NHC(O)NH; where R is hydrogen or lower alkyl and R6 is
 FT lower alkyl"

FT Modified-site 1

FT /note= "Preferably N-terminus is selected from: -NRRL; -
 FT NRC(O)R; -NRC(O)OR; -NRS(O)2R; -NHC(O)NHR; succinimide;
 FT benzyloxycarbonyl-NH; benzyloxycarbonyl-NH with 1-3
 FT substitutions on the phenyl ring selected from lower
 FT alkyl, lower alkoxy, chloro, bromo; where R and R1 are
 FT independently selected from hydrogen and lower alkyl"

FT Modified-site 18

FT /note= "Preferably C-terminus is -C(O)R2 where R2 is
 FT selected from hydroxy, lower alkoxy, and -NR3R4, where R3
 FT and R4 are independently selected from hydrogen and lower
 FT alkyl, and where the nitrogen atom of the -NR3R4 group
 FT can optionally be the amine group of the N-terminus of
 FT the peptide forming a cyclic peptide"

XX WO9640189-A1.

XX PD 19-DEC-1996.

XX PF 05-JUN-1996; 96WO-US008998.

XX PR 07-JUN-1995; 95US-00472371.

XX PR 07-JUN-1995; 95US-00473604.

XX PR 07-JUN-1995; 95US-00476168.

XX PR 07-JUN-1995; 95US-00478128.

XX PR 07-JUN-1995; 95US-00484090.

XX PR 07-JUN-1995; 95US-00485301.

XX (GLAX) GLAXO GROUP LTD.

XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;

CC investigate the mechanism of thrombopoietin signal transduction and
CC receptor activation, or to maintain the proliferation and growth of
CC thrombopoietin dependent cell lines
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

QY 1 TIKGPTLRQWLKSRHTS 18
| | | | | | | | | | | | | | | | | |
DB 1 TIKGPTLRQWLKSRHTS 18

RESULT 4
AAW33026
ID AAW33026 standard; peptide; 18 AA.
XX
AC AAW33026;
XX
XX 11-MAR-1998 (first entry)
DT
DE Thrombopoietin receptor binding peptide.
XX
KW Thrombopoietin receptor; binding peptide; treatment; agonist;
KW haematological disorder; thrombocytopaenia; chemotherapy;
KW radiation therapy; bone marrow transfusion; diagnosis;
KW signal transduction; receptor activation; cell culture.
XX
OS Synthetic.
XX
XX WO9640750-A1.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009623.
XX
XX 07-JUN-1995; 95US-00478128.
XX 07-JUN-1995; 95US-00485301.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barret RW, Cwiria SE, Duffin DJ, Gates CM, Johnson SS;
XX Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-052226/05.
XX
XX Peptides and peptide mimetics which bind to and activate the
XX thrombopoietin receptor - useful in treatment of haematological
XX disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
XX
XX Claim 19; Page 89; 106pp; English.
XX
XX The present peptide binds the thrombopoietin receptor (TR), has a
XX molecular weight of less than 8000 Da and a TR binding affinity as
XX expressed by an IC50 of no more than about 100 microM. It can be used to
XX treat disorders which are susceptible to treatment with a thrombopoietin
XX agonist, preferably haematological disorders and thrombocytopaenia
XX resulting from chemotherapy, radiation therapy or bone marrow
XX transfusions. It can also be used diagnostically, e.g. to investigate the
XX mechanism of thrombopoietin signal transduction and receptor activation,
XX or to maintain the proliferation and growth of thrombopoietin dependent
XX cell lines
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

QY 1 TIKGPTLRQWLKSRHTS 18
| | | | | | | | | | | | | | | | | |

```

Db      1  TIKGPTLQWLKSRHTS 18

RESULT 5
AAAB17024
ID  AAAB17024 standard; peptide; 18 AA.
XX
XX
AC  AAAB17024;
XX
XX  31-OCT-2000 (first entry)
DT
XX
XX  TPO-mimetic peptide sequence SEQ ID NO:80.
DE
XX
XX  Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW  autoimmune disease; cycostatic; antiasthmatic; thrombolytic; VEGF;
KW  immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW  inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW  cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW  vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW  thrombosis; pharmaceutical.
XX
OS  Synthetic.
XX
XX  WO200024782-A2.
XX
XX  04-MAY-2000.
XX
XX  25-OCT-1999; 99WO-US025044.
XX
XX  23-OCT-1998; 98US-0105371P.
XX  22-OCT-1999; 99US-00428082.
XX
XX  (AMGE-) AMGEN INC.
XX
XX  Feige U, Liu C, Cheetham J, Boone TC;
XX
XX  WPI; 2000-350702/30.
XX
XX  Novel composition of matter comprising an Fc domain and pharmacologically
PT  active peptides, useful for treating cancer and autoimmune diseases.
XX
XX  Claim 19; Page 222; 608pp; English.
XX
XX  The present invention describes composition of matter (I) comprising an
CC  Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC  (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC  independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC  (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC  P3, and P4 = are each independently sequences of pharmacologically active
CC  peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC  c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC  of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC  thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC  cells from the present invention can be used for producing pharmaceutical
CC  compositions. The compositions are useful for treating cancer, asthma,
CC  thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC  a Fab domain) can provide a longer half-life or incorporate functions
CC  such as Fc receptor binding, protein A binding, complement fixation, and
CC  possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC  AAB18003 represent nucleotide and amino acid sequences used in the
CC  exemplification of the present invention
XX
XX  Sequence 18 AA;
XX
Query Match      100.0%; Score 97; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TIKGPTLQWLKSRHTS 18
        |||||
Db      1  TIKGPTLQWLKSRHTS 18

```

```

RESULT 6
AAU25869
ID  AAU25869 standard; peptide; 18 AA.
XX
XX
AC  AAU25869;
XX
XX  17-DEC-2001 (first entry)
DT
XX
XX  Human thrombopoietin receptor (TPO-R) activator peptide #55.
DE
XX
XX  Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW  haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW  bone marrow transplantation; haematological disorder; platelet disorder;
KW  enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW  tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW  in vitro expansion; megakaryocyte; Headpiece Dimer gene; laci gene.
XX
OS  Homo sapiens.
XX
XX  US6251864-B1.
XX
XX  26-JUN-2001.
XX
XX  01-MAR-2000; 2000US-00516704.
XX
XX  07-JUN-1995; 95US-00478128.
XX  07-JUN-1995; 95US-00485301.
XX  07-JUN-1996; 96WO-US009623.
XX  15-AUG-1996; 96US-00699027.
XX
XX  (GLAX ) GLAXO GROUP LTD.
XX
XX  Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;
XX  Balasubramanian P, Wagstrom CR, Hendren RW, Poddaturi S;
XX  Yin Q;
XX
XX  WPI; 2001-564142/63.
XX
XX  Activating thrombopoietin receptors in cells, used to treat
PT  thrombocytopenia and hematological disorders, comprises contacting cells
PT  with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX  Disclosure; Col 20; 128pp; English.
XX
XX  Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC  bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC  of activating thrombopoietin receptors in cells comprise contacting the
CC  cells with effective amounts of peptides and peptide mimetics attached to
CC  hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC  as that due to chemotherapy, radiation therapy or bone-marrow
CC  transplantation and to prevent thrombocytopenia in patients at risk. The
CC  sequences are used to treat and prevent haematological disorders
CC  including thrombocytopenia and platelet disorders. They are used in vitro
CC  as unique tools for understanding the biological role of thrombopoietin
CC  (TPO) and to develop other compounds that bind to and activate the TPO
CC  receptor. The peptides can be used to detect TPO receptors on living
CC  cells and fixed cells, in biological fluids, in tissue homogenates, and
CC  in purified or natural biological materials. They may also be used for in
CC  situ staining, fluorescence-activated cell sorting, Western blotting and
CC  enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC  be used for in vitro expansion of megakaryocytes and their committed
CC  progenitors alone or in conjunction with additional cytokines
XX
XX  Sequence 18 AA;
XX
Query Match      100.0%; Score 97; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TIKGPTLQWLKSRHTS 18
        |||||
Db      1  TIKGPTLQWLKSRHTS 18

```


RESULT 7
 AAU25823
 ID AAU25823 standard; peptide; 18 AA.
 XX AC AAU25823;
 XX DT 17-DEC-2001 (first entry)
 XX DE Human thrombopoietin receptor (TPO-R) activator peptide #9.
 XX KW Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
 KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX OS Homo sapiens.
 XX KW US6251864-B1.
 XX PD 26-JUN-2001.
 XX PF 01-MAR-2000; 2000US-00516704.
 XX PR 07-JUN-1995; 95US-00478128.
 XX PR 07-JUN-1995; 95US-00485301.
 XX PR 07-JUN-1996; 96WO-US009623.
 XX PR 15-AUG-1996; 96US-00699027.
 XX PA (GLAXO) GLAXO GROUP LTD.
 XX KW Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ,
 PI Balasubramanian P, Wagstrom CR, Hendren RW, Podduturi S,
 PI Yin Q;
 XX WPI; 2001-564142/63.
 XX DR Activating thrombopoietin receptors in cells, used to treat
 XX thrombocytopenia and hematological disorders, comprises contacting cells
 XX with peptides and peptide mimetics attached to hydrophilic polymers.
 XX FS Disclosure; Col 67-68; 128pp; English.
 XX CC Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
 CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
 CC of activating thrombopoietin receptors in cells comprise contacting the
 CC cells with effective amounts of peptides and peptide mimetics attached to
 CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
 CC as that due to chemotherapy, radiation therapy or bone-marrow
 CC transplantation and to prevent thrombocytopenia in patients at risk. The
 CC sequences are used to treat and prevent haematological disorders
 CC including thrombocytopenia and platelet disorders. They are used in vitro
 CC as unique tools for understanding the biological role of thrombopoietin
 CC (TPO) and to develop other compounds that bind to and activate the TPO
 CC receptor. The peptides can be used to detect TPO receptors on living
 CC cells and fixed cells, in biological fluids, in tissue homogenates, and
 CC in purified or natural biological materials. They may also be used for in
 CC situ staining, fluorescence-activated cell sorting, Western blotting and
 CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
 CC be used for in vitro expansion of megakaryocytes and their committed
 CC progenitors alone or in conjunction with additional cytokines
 XX SQ Sequence 18 AA;
 Query Match 100.0%; Score 97; DB 4; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGPTLRQLKSRHETS 18
 DB 1 TTGPTLRQLKSRHETS 18

RESULT 8
 ABB72910
 ID ABB72910 standard; peptide; 18 AA.
 XX AC ABB72910;
 XX DT 05-APR-2002 (first entry)
 XX DE TPO mimetic peptide SEQ ID NO:80.
 XX KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
 KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TNP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN W0200183525-A2.
 XX PD 08-NOV-2001.
 XX PF 02-MAY-2001; 2001WO-US014310.
 XX PR 03-MAY-2000; 2000US-00563286.
 XX PA (AMGE-) AMGEN INC.
 XX PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX WPI; 2002-130313/17.
 XX DR Novel vehicle-peptide molecule or its multimers useful for treating
 XX inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 XX diabetic retinopathy, obesity, sleep disorders and infertility.
 XX PS Claim 39; Page 44; 176pp; English.
 XX CC The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
 CC neuroprotective activities. (I) can be used as a therapeutic or
 CC prophylactic agent as well as for screening purposes. (I) is useful for
 CC diagnosing diseases characterised by dysfunction of their associated
 CC protein of interest, for identifying normal or abnormal proteins of
 CC interest, as a part of diagnostic kit to detect the presence of their
 CC proteins of interest in a biological sample. Additionally, (I) is useful
 CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
 CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, EPO-
 CC infertility, and neurological degenerative diseases. (I), comprising EPO-
 CC mimetic compounds are useful for treating disorders characterised by low
 CC red blood cell levels such as anaemia. The TPO-mimetic comprising
 CC compounds are useful for treating conditions that involve an existing
 CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
 CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
 CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
 CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
 CC represent amino acid and nucleic acid sequences used in the
 CC exemplification of the present invention
 XX SQ Sequence 18 AA;

Query Match 100.0%; Score 97; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQLKSRHTS 18
| | | | | | | | | | | | | | | | | |
Db 1 TIKGPTLRQLKSRHTS 18

RESULT 9

ADJ73062
ID ADJ73062 standard; peptide; 18 AA.
XX AC ADJ73062;
XX
DT 06-MAY-2004 (first entry)
DE TPO mimetic peptide sequence SeqID 516.
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX Synthetic.
XX WO2003084477-A2.
XX 16-OCT-2003.
XX 24-MAR-2003; 2003WO-US009139.
XX 29-MAR-2002; 2002US-0368791P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Scallion BJ, Ghrayeb J;
XX WPI; 2003-804237/75.
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.

XX Disclosure; SEQ ID NO 516; 97pp; English.

XX This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 97; DB 7; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQLKSRHTS 18
| | | | | | | | | | | | | | | | | |
Db 1 TIKGPTLRQLKSRHTS 18

RESULT 10

ADJ52697
ID ADJ52697 standard; peptide; 18 AA.
XX AC ADJ52697;
XX
DT 06-MAY-2004 (first entry)
DE CH1 deleted mimetibody-related peptide SeqID516.
XX CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX Unidentified.
OS Synthetic.
XX WO2004002417-A2.
XX 08-JAN-2004.
XX 27-JUN-2003; 2003WO-US020347.
XX 28-JUN-2002; 2002US-0392431P.
XX (CENZ) CENTOCOR INC.
XX Heavner GA, Knight DM, Ghrayeb J, Scallion BJ, Nesspor TC;
XX Kutoloski RA;
XX WPI; 2004-082870/08.
XX New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX Claim 2; SEQ ID NO 516; 129pp; English.

XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating the symptoms of an
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX Sequence 18 AA;

Query Match 100.0%; Score 97; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TIKGPTLRQLKSRHTS 18
| | | | | | | | | | | | | | | | | |
Db 1 TIKGPTLRQLKSRHTS 18

RESULT 11

ADJ51658
ID ADJ51658 standard; peptide; 18 AA.
XX AC
XX ADJ51658;
XX DT
XX 06-MAY-2004 (first entry)
XX DE
XX CHI deleted mimetibody-related peptide SeqIDS16.
XX KW CHI deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX OS Unidentified.
OS OS Synthetic.
XX PN WO2004002424-A2.
XX XX
XX PD 08-JAN-2004.
XX PF 30-JUN-2003; 2003WO-US020495.
XX PR 28-JUN-2002; 2002US-0392431P.
XX PR 19-SEP-2002; 2002US-0412144P.
XX PA (CENZ) CENTOCOR INC.
XX XX
XX XX Heaven GA, Knight DM, Ghayeb J, Scallon BJ, Nesspor TC;
PI Kutloski KA;
XX WPI; 2004-082872/08.
XX
XX New CHI deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 15; SEQ ID NO 516; 123pp; English.
XX
XX This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cycostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CHI deleted mimetibody, such as a bone or joint,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
XX mimetibody of the invention.
XX
XX Sequence 18 AA;
SQ

Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TIKGPTLRQWLKSRHTS 18
DB 1 TIKGPTLRQWLKSRHTS 18
RESULT 12
AAW09491
ID AAW09491 standard; protein; 19 AA.
XX AC AAW09491;
XX DT 10-SEP-1997 (first entry)
XX DE Thrombopoietin receptor binding peptide.
XX KW Haematology; thrombocytopenia; TPO; TR; proliferation;
KW bone marrow transfusion; chemotherapy; radiation therapy.
XX OS Synthetic.
XX PN WO9640189-A1.
XX PD 19-DEC-1996.
XX PF 05-JUN-1996; 96WO-US008998.
XX PR 07-JUN-1995; 95US-00472371.
XX PR 07-JUN-1995; 95US-00473604.
XX PR 07-JUN-1995; 95US-00476168.
XX PR 07-JUN-1995; 95US-00478128.
XX PR 07-JUN-1995; 95US-00484090.
XX PR 07-JUN-1995; 95US-00485301.
XX PA (GLAX) GLAXO GROUP LTD.
XX XX
XX Dower WJ, Barrett RW, Cwirila SE, Duffin DJ, Gates CM, Johnson SS;
PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
XX WPI; 1997-051883/05.
XX
XX Thrombopoietin receptor-binding/activating peptide(s) and peptide
PT mimetic(s) - useful in treatment of haematological disorders, esp.
PT thrombocytopenia resulting from chemotherapy, etc.
XX
XX Disclosure; Page 26; 106pp; English.
XX
XX The present sequence is a peptide which binds to thrombopoietin (TPO)
CC receptor (TR). The compound can be used for treating patients suffering
CC from haematological disorders and thrombocytopenia resulting from
CC chemotherapy, radiation therapy or bone marrow transfusions. The peptide
CC may also be used to maintain the proliferation and growth of TPO-
CC dependent cell lines and for use in biological research, for detecting
CC TPO receptors on living cells
XX
XX Sequence 19 AA;
SQ

Query Match 64.9%; Score 63; DB 2; Length 19;
Best Local Similarity 73.3%; Pred. No. 0.0059;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 GTPLRQWLKSRHTS 18
DB 5 GPTLRQWLAAARNHLS 19
RESULT 13
AAW35418
ID AAW35418 standard; peptide; 19 AA.
XX AC AAW35418;
XX

XX 11-MAR-1998 (first entry)
 XX Thrombopoietin receptor binding peptide.
 DE
 XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 XX haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Cross-links 3
 FT /note= "linked via disulfide bond to Cys3 of identical
 FT peptide"
 FT Modified-site 19
 FT /note= "NH2-Ser"
 XX
 XX WO9640750-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 07-JUN-1996; 96WO-US009623.
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR
 XX 07-JUN-1995; 95US-00485301.
 XX
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 PI
 XX WPI; 1997-052226/05.
 DR
 XX Peptides and peptide mimetics which bind to and activate the
 XX thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
 PT
 XX Disclosure; Page 26; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 XX Sequence 19 AA;
 SQ
 Query Match 64.9%; Score 63; DB 2; Length 19;
 Best Local Similarity 73.3%; Pred. No. 0.0059;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GPTLRQWLKSRHTS 18
 DB 5 GPTLRQWLARNHLS 19
 ||||| :|||
 RESULT 14
 AAW36642
 ID AAW36642 standard; peptide; 19 AA.
 XX
 XX AAW36642;
 AC
 XX 11-MAR-1998 (first entry)
 DT Thrombopoietin receptor binding peptide.
 XX
 XX Thrombopoietin receptor; binding peptide; treatment; agonist;
 XX haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Cross-links 3
 FT /note= "linked via disulfide bond to Cys3 of identical
 FT peptide"
 FT Modified-site 19
 FT /note= "NH2-Ser"
 XX
 XX WO9640750-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 07-JUN-1996; 96WO-US009623.
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR
 XX 07-JUN-1995; 95US-00485301.
 XX
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 PI
 XX WPI; 1997-052226/05.
 DR
 XX Peptides and peptide mimetics which bind to and activate the
 XX thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
 PT
 XX Disclosure; Page 73; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 XX Sequence 19 AA;
 SQ
 Query Match 64.9%; Score 63; DB 2; Length 19;
 Best Local Similarity 73.3%; Pred. No. 0.0059;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GPTLRQWLKSRHTS 18
 DB 5 GPTLRQWLARNHLS 19
 ||||| :|||
 RESULT 15
 AAU25861
 ID AAU25861 standard; peptide; 19 AA.
 XX
 XX AAU25861;
 AC
 XX 17-DEC-2001 (first entry)
 DT Human thrombopoietin receptor (TPO-R) activator peptide #47.
 XX
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopaenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX
 XX Homo sapiens.
 OS
 XX US6251864-B1.
 PN
 XX 26-JUN-2001.
 PD
 XX 01-MAR-2000; 2000US-00516704.
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR

KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX Synthetic.
 OS
 XX WO9640750-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 07-JUN-1996; 96WO-US009623.
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR
 XX 07-JUN-1995; 95US-00485301.
 XX
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX Dower WJ, Barret RW, Cwirla SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 PI
 XX WPI; 1997-052226/05.
 DR
 XX Peptides and peptide mimetics which bind to and activate the
 XX thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
 PT
 XX Disclosure; Page 26; 106pp; English.
 PS
 XX The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 XX Sequence 19 AA;
 SQ
 Query Match 64.9%; Score 63; DB 2; Length 19;
 Best Local Similarity 73.3%; Pred. No. 0.0059;
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 GPTLRQWLKSRHTS 18
 DB 5 GPTLRQWLARNHLS 19
 ||||| :|||
 RESULT 15
 AAU25861
 ID AAU25861 standard; peptide; 19 AA.
 XX
 XX AAU25861;
 AC
 XX 17-DEC-2001 (first entry)
 DT Human thrombopoietin receptor (TPO-R) activator peptide #47.
 XX
 XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
 KW haemostatic; thrombocytopaenia; chemotherapy; radiation therapy; ELISA;
 KW bone marrow transplantation; haematological disorder; platelet disorder;
 KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
 KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
 KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
 XX
 XX Homo sapiens.
 OS
 XX US6251864-B1.
 PN
 XX 26-JUN-2001.
 PD
 XX 01-MAR-2000; 2000US-00516704.
 XX
 XX 07-JUN-1995; 95US-00478128.
 PR


```

PF 17-NOV-2003; 2003WO-US036894.
XX
PR 02-DEC-2002; 2002US-00307724.
XX
XX (ALEX-) ALEXION PHARM INC.
XX
XX Bowdish KS, Frederickson S, Renshaw M;
XX
XX WPI; 2004-460973/43.
XX
XX DR N-PSDB; ADQ16626.
XX
XX New immunoglobulin molecule comprising a region, where two
XX complementarity determining regions (CDRs) are replaced with EPO mimetic
XX or a TPO mimetic, useful for treating thrombocytopenia.
XX
XX Example 1; SEQ ID NO 45; 107pp; English.
XX
XX The invention relates to a novel immunoglobulin molecule or its fragment
XX comprising a region where amino acid residues corresponding to at least a
XX portion of a two complementarity determining regions (CDRs) are replaced
XX with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
XX a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
XX invention has immunosuppressive activity, and may have a use in
XX immunotherapy. The immunoglobulin molecule is useful for diagnosing or
XX treating thrombocytopenia as a result of chemotherapy, bone marrow
XX transplantation, or chronic diseases such as idiopathic thrombocytopenia.
XX The present sequence represents a TPO mimetic peptide with flanking
XX residues.
XX
XX Sequence 18 AA;
XX
XX Query Match 63.9%; Score 62; DB 8; Length 18;
XX Best Local Similarity 78.6%; Pred. NO. 0.0081;
XX Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 TIKGPTLRQWLKSR 14
XX ||:|||||:|
XX Db 2 TIEGPTLRQWLAAAR 15
XX ||:|||||:|
XX
XX RESULT 21
XX ADV44363
XX ID ADV44363 standard; peptide; 18 AA.
XX
XX AC ADV44363;
XX
XX XX
XX DT 10-MAR-2005 (first entry)
XX
XX XX
XX DE Agonist TPO mimetic peptide seqid 45.
XX
XX XX anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
XX KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
XX KW transplant rejection; prophylaxis; myeloproliferative disorder;
XX KW hematological disease; TPO mimetic peptide.
XX
XX OS Unidentified.
XX
XX XX
XX PN WO2004108078-A2.
XX
XX XX
XX PD 16-DEC-2004.
XX
XX XX
XX PF 26-MAY-2004; 2004WO-US016574.
XX
XX XX
XX PR 02-JUN-2003; 2003US-00452590.
XX
XX XX (ALEX-) ALEXION PHARM INC.
XX
XX XX Bowdish KS, Frederickson S, Renshaw M, Orenca C;
XX
XX XX WPI; 2005-031588/03.
XX
XX DR N-PSDB; ADV44364.
XX
XX XX
XX PT New immunoglobulin molecule comprises a region where amino acid residues
XX corresponding to a portion of complementarity determining region (CDR) is
XX replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
XX patients.
XX
XX XX
XX PS Example 1; SEQ ID NO 45; 139pp; English.
XX
XX The invention describes an immunoglobulin molecule or its fragment
XX comprising a region where amino acid residues corresponding to at least
XX a portion of two CDRs are replaced with a peptide mimetic selected from
XX an EPO mimetic or a TPO mimetic; or a region where amino acid residues
XX corresponding to at least a portion of a CDR is replaced by a peptide
XX mimetic including SEQ ID NO. 126 (not defined in the specification),
XX where X at each occurrence represents any amino acid. Also described are:

```

CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
 CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an
 CC immunoglobulin molecule or its fragment; and a pharmaceutical carrier. Also
 CC disclosed are: engineering immunoglobulin molecules or fragments;
 CC creation of a library of monoclonal antibodies; stimulating
 CC proliferation, differentiation, or growth of megakaryocytes; and
 CC activating a homodimeric receptor protein. The immunoglobulin molecules
 CC are useful for treating HIV-infected patients, patients undergoing
 CC chemotherapy, bone marrow transplant patients, stem cell transplant
 CC patients, or patients suffering from myeloproliferative disorders. This
 CC is the amino acid sequence of an agonist TPO mimetic peptide used to
 CC replace a CDR in a rationally designed antibody.

XX
 SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
 Best Local Similarity 78.6%; Pred. No. 0.0081;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLQWLKSR 14
 ||:|||||||:
 DB 2 TIEGPTLQWLAAR 15

RESULT 22

ADV44367
 ID ADV44367 standard; peptide; 18 AA.

XX
 AC ADV44367;

DT 10-MAR-2005 (first entry)

XX Agonist TPO mimetic peptide seqid 49.

XX anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
 KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
 KW transplant rejection; prophyaxis; myeloproliferative disorder;
 KW hematological disease; TPO mimetic peptide.

XX Unidentified.

XX WO2004108078-A2.

XX 16-DEC-2004.

XX 26-MAY-2004; 2004WO-US016574.

XX 02-JUN-2003; 2003US-00452590.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M, Orenicia C;

XX WPI; 2005-031588/03.

XX N-PSDB; ADV44368.

XX New immunoglobulin molecule comprises a region where amino acid residues
 PT corresponding to a portion of complementarity determining region (CDR) is
 PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
 PT patients.

XX Example 1; SEQ ID NO 49; 139pp; English.

XX The invention describes an immunoglobulin molecule or its fragment
 CC comprising: a region where amino acid residues corresponding to at least
 CC a portion of two CDRs are replaced with a peptide mimetic selected from
 CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced by a peptide
 CC mimetic including SEQ ID NO. 126 (not defined in the specification),
 CC where X at each occurrence represents any amino acid. Also described are:
 CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an

CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an
 CC immunoglobulin molecule or its fragment; and a pharmaceutical carrier. Also
 CC disclosed are: engineering immunoglobulin molecules or fragments;
 CC creation of a library of monoclonal antibodies; stimulating
 CC proliferation, differentiation, or growth of megakaryocytes; and
 CC activating a homodimeric receptor protein. The immunoglobulin molecules
 CC are useful for treating HIV-infected patients, patients undergoing
 CC chemotherapy, bone marrow transplant patients, stem cell transplant
 CC patients, or patients suffering from myeloproliferative disorders. This
 CC is the amino acid sequence of an agonist TPO mimetic peptide used to
 CC replace a CDR in a rationally designed antibody.

XX
 SQ Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
 Best Local Similarity 78.6%; Pred. No. 0.0081;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLQWLKSR 14
 ||:|||||||:
 DB 2 TIEGPTLQWLAAR 15

RESULT 23

AE12840
 ID AEB12840 standard; peptide; 18 AA.

XX
 AC AEB12840;

DT 08-SEP-2005 (first entry)

XX TPO mimetic peptide, X7c.

XX TPO; Thrombopoietin; protein therapy; antibody engineering;
 KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
 KW cardiac failure; diabetes; obesity.

XX Synthetic.

XX WO2005060642-A2.

XX 07-JUL-2005.

XX 15-DEC-2004; 2004WO-US041946.

XX 15-DEC-2003; 2003US-00737290.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M, Orenicia C;

XX WPI; 2005-479402/48.

XX N-PSDB; AEB12841.

XX New immunoglobulin molecule comprising a region where amino acid residues
 PT corresponding to at least a portion of a complementarity determining
 PT region is replaced with a peptide, for treating congestive heart failure,
 PT diabetes or obesity.

XX Example 1; SEQ ID NO 49; 152pp; English.

XX The invention relates to an immunoglobulin (Ig) molecule or its fragment
 CC comprising a region where amino acid residues corresponding to at least a
 CC portion of a complementarity determining regions (CDR) is replaced with a
 CC peptide selected from human brain natriuretic protein (hBNP), hBNP
 CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
 CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
 CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
 CC molecule, an expression vector comprising the nucleic acid, a host cell
 CC transformed with the expression vector, producing an immunoglobulin
 CC molecule (or its fragment, comprising culturing the host cell under

CC conditions suitable for expression of the immunoglobulin or its
 CC fragment), a composition comprising an immunoglobulin (or its fragment)
 CC and a pharmaceutically acceptable carrier, treating congestive heart
 CC failure (comprising administering to the subject an immunoglobulin
 CC molecule or fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with a peptide
 CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
 CC (comprising administering to a subject an immunoglobulin molecule or its
 CC fragment comprising a region where amino acid residues corresponding to
 CC at least a portion of a CDR is replaced with a peptide selected from GLP-
 CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
 CC glucagons, glucagons mimetics and PACAP-38), preserving/improving beta-
 CC cell function (comprising administering to a subject an immunoglobulin
 CC molecule or fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with GLP-1),
 CC inducing endothelial-dependent relaxation of precontracted pulmonary
 CC artery rings (comprising administering to a subject an immunoglobulin
 CC molecule or fragment thereof comprising a region where amino acid
 CC residues corresponding to at least a portion of a CDR is replaced with
 CC GLP-1) and administering to a subject an immunoglobulin molecule or its
 CC fragment (comprising a region where amino acid residues corresponding to
 CC at least a portion of a complementarity determining regions (CDR) is
 CC replaced with a thiazolidinedione derivative), regulating adiponectin
 CC expression (comprising administering to a subject an immunoglobulin
 CC molecule or its fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with a
 CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
 CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
 CC partially replaced with a peptide listed above or (as described in the
 CC examples) a Thrombopoietin (TPO) mimetic, erythropoietin (EPO) mimetic
 CC or ANP (atrial natriuretic peptide). The molecule, composition and
 CC methods are useful for treating congestive heart failure, diabetes or
 CC obesity. The present sequence is a TPO mimetic peptide clone flanked by
 CC random amino acids for inclusion in an immunoglobulin of the invention.

XX Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
 Best Local Similarity 78.6%; Pred. No. 0.0081;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TIKGPTLRQLKSR 14
 DB 2 TIEGPTLRQLAAR 15

RESULT 24
 AEB12836
 ID AEB12836 standard; peptide; 18 AA.

XX AC AEB12836;

XX 08-SEP-2005 (first entry)

DE TPO mimetic peptide, X5C.

XX TPO; Thrombopoietin; protein therapy; antibody engineering;
 KW hematopoiesis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
 KW cardiac failure; diabetes; obesity.

XX Synthetic.

XX WO2005060642-A2.

XX 07-JUL-2005.

XX 15-DEC-2004; 2004WO-US041946.

XX 15-DEC-2003; 2003US-00737290.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Frederickson S, Renshaw M, Orencia C;

XX

DR WPI; 2005-479402/48.

DR N-PSDB; AEB12837.

XX New immunoglobulin molecule comprising a region where amino acid residues
 PT corresponding to at least a portion of a complementarity determining
 PT region is replaced with a peptide, for treating congestive heart failure,
 PT diabetes or obesity.

XX Example 1; SEQ ID NO 45; 152pp; English.

XX The invention relates to an immunoglobulin (Ig) molecule or its fragment
 CC comprising a region where amino acid residues corresponding to at least a
 CC portion of a complementarity determining regions (CDR) is replaced with a
 CC peptide selected from human brain natriuretic protein (hBNP), hBNP
 CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
 CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
 CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
 CC molecule, an expression vector comprising the nucleic acid, a host cell
 CC transformed with the expression vector, producing an immunoglobulin
 CC molecule (or its fragment, comprising culturing the host cell under
 CC conditions suitable for expression of the immunoglobulin or its
 CC fragment), a composition comprising an immunoglobulin (or its fragment)
 CC and a pharmaceutically acceptable carrier, treating congestive heart
 CC failure (comprising administering to the subject an immunoglobulin
 CC molecule or fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with a peptide
 CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
 CC (comprising administering to a subject an immunoglobulin molecule or its
 CC fragment comprising a region where amino acid residues corresponding to
 CC at least a portion of a CDR is replaced with a peptide selected from GLP-
 CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
 CC glucagons, glucagons mimetics and PACAP-38), preserving/improving beta-
 CC cell function (comprising administering to a subject an immunoglobulin
 CC molecule or fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with GLP-1),
 CC inducing endothelial-dependent relaxation of precontracted pulmonary
 CC artery rings (comprising administering to a subject an immunoglobulin
 CC molecule or fragment thereof comprising a region where amino acid
 CC residues corresponding to at least a portion of a CDR is replaced with
 CC GLP-1) and administering to a subject an immunoglobulin molecule or its
 CC fragment (comprising a region where amino acid residues corresponding to
 CC at least a portion of a complementarity determining regions (CDR) is
 CC replaced with a thiazolidinedione derivative), regulating adiponectin
 CC expression (comprising administering to a subject an immunoglobulin
 CC molecule or its fragment comprising a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced with a
 CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
 CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
 CC partially replaced with a peptide listed above or (as described in the
 CC examples) a Thrombopoietin (TPO) mimetic, erythropoietin (EPO) mimetic
 CC or ANP (atrial natriuretic peptide). The molecule, composition and
 CC methods are useful for treating congestive heart failure, diabetes or
 CC obesity. The present sequence is a TPO mimetic peptide clone flanked by
 CC random amino acids for inclusion in an immunoglobulin of the invention.

XX Sequence 18 AA;

Query Match 63.9%; Score 62; DB 9; Length 18;
 Best Local Similarity 78.6%; Pred. No. 0.0081;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TIKGPTLRQLKSR 14

DB 2 TIEGPTLRQLAAR 15

RESULT 25

AAW09493

ID AAW09493 standard; protein; 19 AA.

XX AAW09493;


```

PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX WPI; 2001-564142/63.
XX
XX Activating thrombopoietin receptors in cells, used to treat
XX thrombocytopenia and hematological disorders, comprises contacting cells
XX with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 20; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
XX bind to and activate the human thrombopoietin receptor (TPO-R). Methods
XX of activating thrombopoietin receptors in cells comprise contacting the
XX cells with effective amounts of peptides and peptide mimetics attached to
XX hydrophilic polymers. The methods are used to treat thrombocytopenia such
XX as that due to chemotherapy, radiation therapy or bone-marrow
XX transplantation and to prevent thrombocytopenia in patients at risk. The
XX sequences are used to treat and prevent hematological disorders
XX including thrombocytopenia and platelet disorders. They are used in vitro
XX as unique tools for understanding the biological role of thrombopoietin
XX (TPO) and to develop other compounds that bind to and activate the TPO
XX receptor. The peptides can be used to detect TPO receptors on living
XX cells and fixed cells, in biological fluids, in tissue homogenates, and
XX in purified or natural biological materials. They may also be used for in
XX situ staining, fluorescence-activated cell sorting, Western blotting and
XX enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
XX be used for in vitro expansion of megakaryocytes and their committed
XX progenitors alone or in conjunction with additional cytokines
XX
XX Sequence 19 AA;
XX
XX Query Match 63.9%; Score 62; DB 4; Length 19;
XX Best Local Similarity 73.3%; Pred. No. 0.0086;
XX Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
QY 4 GPTLRQWLKSRHTS 18
DB ||||| :|||
5 GPTLRQWLKSRHTS 19

RESULT 28
ABG1748
ID ABG1748 standard; protein; 144 AA.
XX
XX ABG1748;
XX
XX 20-JAN-2003 (first entry)
XX
XX Antibody CDR containing MPL-TPO binding sequence, TPOVHCDR1.
XX
XX Agonist; immunoglobulin; Ig; variable domain; heavy chain; light chain;
XX complementarity determining region; CDR; antigenic; thrombopoietin; TPO;
XX thrombopoietin receptor; MPL; cytotoxic T-lymphocyte; CTL; epitope;
XX T-helper cell; B-helper cell; synthebody; pharmaceutical; vaccine;
XX proliferation; growth; differentiation; haematopoietic cell; antibody;
XX platelet progenitor cell; immune disorder; thrombocytopenia;
XX disseminated intravascular coagulation; stem cell; transplantation;
XX gene therapy; diagnostic; haemostatic; immunomodulator; anticoagulant;
XX consensus variable heavy chain domain; CONVH.
XX
XX Synthetic.
XX Unidentified.
XX
XX Key Location/Qualifiers
XX Region 50..63
XX /note= "TPO receptor binding agonist peptide"
XX
XX WO200278612-A2.
XX
XX 10-OCT-2002.
XX
XX 02-APR-2002; 2002WO-US010301.
XX PF

02-APR-2001; 2001US-0281183P.
(PURD ) PURDUE PHARMA LP.
Soltis DA, Burch RM, Ogert RA;
WPI; 2003-040615/03.
New thrombopoietin synthebodies, useful for stimulating proliferation,
growth, or differentiation of hematopoietic cells, for treating or
preventing hematopoietic or immune disorders, e.g. thrombocytopenia.
Example 1; Page 75; 97pp; English.
The invention discloses a variant of an immunoglobulin (Ig) variable
heavy or light chain domain that comprises at least one complementarity
determining region (CDR) and framework regions flanking the CDR. The CDR
also has added or substituted to it, at least one binding sequence which
is heterologous to the CDR and is an antigenic, agonistic sequence from a
thrombopoietin (TPO) receptor binding sequence. The antigenic sequence
can be a binding sequence heterologous to the CDR, a cytotoxic T-
lymphocyte (CTL)-epitope sequence, a T-helper cell sequence, a B-helper
cell sequence or a combination of each. The variant or thrombopoietin
synthebody, pharmaceutical and vaccine compositions are useful for
stimulating proliferation, growth or differentiation of haematopoietic
cells, particularly platelet progenitor cells. The variants are also
useful for treating or preventing haematopoietic or immune disorders
resulting from chemotherapy, radiation therapy, or bone marrow
transfusions (e.g. thrombocytopenia or disseminated intravascular
coagulation). Compositions comprising the synthebodies can be used for
the mobilisation, amplification and ex vivo expansion of stem cells and
committed precursor cells for autologous and allogeneic transplantation
as well as for the expansion of stem cells for gene therapy. They are
also useful as diagnostic or analytical reagents for studying the
function of thrombopoietin and its receptor in vivo or in vitro. The
sequence presented, TPOVHCDR1, is the TPO receptor (MPL) agonist peptide
sequence contained within the immunoglobulin CDR consensus variable heavy
chain domain (CONVH)
XX Sequence 144 AA;
XX
XX Query Match 63.9%; Score 62; DB 6; Length 144;
XX Best Local Similarity 78.6%; Pred. No. 0.087;
XX Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 TIKGPTLRQWLKSR 14
DB ||||| :|||
49 TIEGPTLRQWLKSR 62

RESULT 29
AAU26006
ID AAU26006 standard; peptide; 14 AA.
XX
XX AAU26006;
XX
XX 17-DEC-2001 (first entry)
XX
XX Human thrombopoietin receptor (TPO-R) activator peptide #192.
XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; cytokine;
XX haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
XX bone marrow transplantation; haematological disorder; platelet disorder;
XX enzyme-linked immunosorbent assay; in situ staining; biological fluid;
XX tissue homogenate; fluorescence-activated cell sorting; Western blotting;
XX in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
XX Homo sapiens.
XX
XX US6251864-B1.
XX
XX 26-JUN-2001.
XX PD

```


XX The present peptide binds the thrombopoietin receptor (TR), has a
 CC molecular weight of less than 8000 Da and a TR binding affinity as
 CC expressed by an IC50 of no more than about 100 microm. It can be used to
 CC treat disorders which are susceptible to treatment with a thrombopoietin
 CC agonist, preferably haematological disorders and thrombocytopaenia
 CC resulting from chemotherapy, radiation therapy or bone marrow
 CC transfusions. It can also be used diagnostically, e.g. to investigate the
 CC mechanism of thrombopoietin signal transduction and receptor activation,
 CC or to maintain the proliferation and growth of thrombopoietin dependent
 CC cell lines
 XX
 SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
 Best Local Similarity 71.4%; Pred. No. 0.035;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
 : : : : : : : : : : : : : : : : : :
 Db 1 SIEGPTLRWLTSR 14
 : : : : : : : : : : : : : : : : : :

RESULT 34
 AAW36652
 ID AAW36652 standard; peptide; 18 AA.
 XX
 AC AAW36652;
 XX
 DT 11-MAR-1998 (first entry)
 XX
 DE Thrombopoietin receptor binding peptide.
 XX
 KW Thrombopoietin receptor; binding peptide; treatment; agonist;
 KW haematological disorder; thrombocytopaenia; chemotherapy;
 KW radiation therapy; bone marrow transfusion; diagnosis;
 KW signal transduction; receptor activation; cell culture.
 XX
 OS Synthetic.
 XX
 PN WO9640750-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US009623.
 XX
 PR 07-JUN-1995; 95US-00478128.
 PR 07-JUN-1995; 95US-00485301.
 XX
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 PI Dower WJ, Barret RW, Cwirle SE, Duffin DJ, Gates CM, Johnson SS;
 PI Mattheakis LC, Schatz PJ, Wagstrom CR, Wrighton NC;
 XX
 DR WPI; 1997-052226/05.
 XX
 XX Peptides and peptide mimetics which bind to and activate the
 PT thrombopoietin receptor - useful in treatment of haematological
 PT disorders, esp. thrombocytopaenia resulting from chemotherapy, etc.
 XX
 PS Disclosure; Page 27; 106pp; English.
 XX
 CC The present peptide, which binds the thrombopoietin receptor (TR), can be
 CC used to treat disorders which are susceptible to treatment with a
 CC thrombopoietin agonist, preferably haematological disorders and
 CC thrombocytopaenia resulting from chemotherapy, radiation therapy or bone
 CC marrow transfusions. It can also be used diagnostically, e.g. to
 CC investigate the mechanism of thrombopoietin signal transduction and
 CC receptor activation, or to maintain the proliferation and growth of
 CC thrombopoietin dependent cell lines
 XX
 SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 2; Length 18;
 Best Local Similarity 71.4%; Pred. No. 0.035;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
 : : : : : : : : : : : : : : : : : :
 Db 1 SIEGPTLRWLTSR 14
 : : : : : : : : : : : : : : : : : :

RESULT 35
 AAB17026
 ID AAB17026 standard; peptide; 18 AA.
 XX
 AC AAB17026;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE TPO-mimetic peptide sequence SEQ ID NO:82.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; BPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
 KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
 KW thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US025044.
 XX
 PR 23-OCT-1998; 98US-0105371P.
 PR 22-OCT-1999; 99US-00428082.
 XX
 PA (AMGEN) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheatham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and pharmacologically
 PT active peptides, useful for treating cancer and autoimmune diseases.
 XX
 PS Claim 19; Page 22; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
 CC (L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
 CC P3, and P4 = are each independently sequences of pharmacologically active
 CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
 CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
 CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
 CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
 CC cells from the present invention can be used for producing pharmaceutical
 CC compositions. The compositions are useful for treating cancer, asthma,
 CC thrombosis, or autoimmune diseases. The use of an Fc domain rather than
 CC a Fab domain) can provide a longer half-life or incorporate functions
 CC such as Fc receptor binding, protein A binding, complement fixation, and
 CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
 CC AAB18003 represent nucleotide and amino acid sequences used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 3; Length 18;
 Best Local Similarity 71.4%; Pred. No. 0.035;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

		Matches	10;	Conservative	3;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	TIKGPTLRQWLKSR	14								
		: : : :									
Db	1	SIEGPTLRWLTSR	14								
RESULT 36											
AAU25868											
ID	AAU25868 standard; peptide; 18 AA.										
XX	AC	AAU25868;									
XX	DT	17-DEC-2001 (first entry)									
XX	DE	Human thrombopoietin receptor (TPO-R) activator peptide #54.									
XX	XX	Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;									
KW	KW	haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;									
KW	KW	bone marrow transplantation; haematological disorder; platelet disorder;									
KW	KW	enzyme-linked immunosorbent assay; in situ staining; biological fluid;									
KW	KW	tissue homogenate; fluorescence-activated cell sorting; Western blotting;									
KW	KW	in vitro expansion; megakaryocyte; Headpiece dimer gene; lacI gene.									
XX	OS	Homo sapiens.									
XX	XX	US6251864-B1.									
XX	XX	26-JUN-2001.									
XX	XX	01-MAR-2000; 2000US-00516704.									
XX	XX	07-JUN-1995; 95US-00478128.									
PR	PR	07-JUN-1995; 95US-00485301.									
PR	PR	07-JUN-1996; 96WO-US009623.									
PR	PR	15-AUG-1996; 96US-00699027.									
XX	XX	(GLAXO) GLAXO GROUP LTD.									
XX	XX	Dower WJ, Barrett RW, Cwirla SE, Gates CM, Schatz PJ;									
PI	PI	Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;									
PI	PI	Yin Q;									
DR	DR	WPI; 2001-564142/63.									
XX	XX	Activating thrombopoietin receptors in cells, used to treat									
PT	PT	thrombocytopenia and hematological disorders, comprises contacting cells									
PT	PT	with peptides and peptide mimetics attached to hydrophilic polymers.									
XX	XX	Disclosure; Col 20; 128pp; English.									
PS	PS	Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that									
XX	XX	bind to and activate the human thrombopoietin receptor (TPO-R). Methods									
CC	CC	of activating thrombopoietin receptors in cells comprise contacting the									
CC	CC	cells with effective amounts of peptides and peptide mimetics attached to									
CC	CC	hydrophilic polymers. The methods are used to treat thrombocytopenia such									
CC	CC	as that due to chemotherapy, radiation therapy or bone-marrow									
CC	CC	transplantation and to prevent thrombocytopenia in patients at risk. The									
CC	CC	sequences are used to treat and prevent haematological disorders									
CC	CC	including thrombocytopenia and platelet disorders. They are used in vitro									
CC	CC	as unique tools for understanding the biological role of thrombopoietin									
CC	CC	(TPO) and to develop other compounds that bind to and activate the TPO									
CC	CC	receptor. The peptides can be used to detect TPO receptors on living									
CC	CC	cells and fixed cells, in biological fluids, in tissue homogenates, and									
CC	CC	in purified or natural biological materials. They may also be used for in									
CC	CC	situ staining, fluorescence-activated cell sorting, Western blotting and									
CC	CC	enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can									
CC	CC	be used for in vitro expansion of megakaryocytes and their committed									
CC	CC	progenitors alone or in conjunction with additional cytokines									
XX	XX	Sequence 18 AA;									
SQ	SQ	Query Match 59.8%; Score 58; DB 4; Length 18;									
		Best Local Similarity 71.4%; Pred. No. 0.035;									

Query Match 59.8%; Score 58; DB 4; Length 18;

Best Local Similarity 71.4%; Pred. No. 0.035; Mismatches 3; Indels 1; Gaps 0; Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
:|||||:|||||
Db 1 SIEGPTLRWLTSR 14

RESULT 38
AAU25871
ID AAU25871 standard; peptide; 18 AA.
XX AAU25871;
DT 17-DEC-2001 (first entry)
XX Human thrombopoietin receptor (TPO-R) activator peptide #57.
DE XX
XX Peptide mimetic; human; thrombopoietin receptor; TPO-R; TPO; cytokine;
KW haemostatic; thrombocytopenia; chemotherapy; radiation therapy; ELISA;
KW bone marrow transplantation; haematological disorder; platelet disorder;
KW enzyme-linked immunosorbent assay; in situ staining; biological fluid;
KW tissue homogenate; fluorescence-activated cell sorting; Western blotting;
KW in vitro expansion; megakaryocyte; Headpiece Dimer gene; lacI gene.
XX
OS Homo sapiens.
XX
PN US6251864-B1.
XX
PD 26-JUN-2001.
XX
XX 01-MAR-2000; 2000US-00516704.
XX
XX 07-JUN-1995; 95US-00478128.
PR 07-JUN-1995; 95US-00485301.
PR 07-JUN-1996; 96WO-US009623.
PR 15-AUG-1996; 96US-00699027.
XX
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Dower WJ, Barrett RW, Cwiria SE, Gates CM, Schatz PJ;
PI Balasubramanian P, Wagstrom CR, Hendren RW, Deprince RB, Podduturi S;
PI Yin Q;
XX
XX WPI; 2001-564142/63.
DR
XX Activating thrombopoietin receptors in cells, used to treat
PT thrombocytopenia and hematological disorders, comprises contacting cells
PT with peptides and peptide mimetics attached to hydrophilic polymers.
XX
XX Disclosure; Col 20; 128pp; English.
XX
XX Sequences AAU25815-AAU26049 represent peptides and peptide mimetics that
CC bind to and activate the human thrombopoietin receptor (TPO-R). Methods
CC of activating thrombopoietin receptors in cells comprise contacting the
CC cells with effective amounts of peptides and peptide mimetics attached to
CC hydrophilic polymers. The methods are used to treat thrombocytopenia such
CC as that due to chemotherapy, radiation therapy or bone-marrow
CC transplantation and to prevent thrombocytopenia in patients at risk. The
CC sequences are used to treat and prevent haematological disorders
CC including thrombocytopenia and platelet disorders. They are used in vitro
CC as unique tools for understanding the biological role of thrombopoietin
CC (TPO) and to develop other compounds that bind to and activate the TPO
CC receptor. The peptides can be used to detect TPO receptors on living
CC cells and fixed cells, in biological fluids, in tissue homogenates, and
CC in purified or natural biological materials. They may also be used for in
CC situ staining, fluorescence-activated cell sorting, Western blotting and
CC enzyme-linked immunosorbent assay (ELISA). In addition, the peptides can
CC be used for in vitro expansion of megakaryocytes and their committed
XX progenitors alone or in conjunction with additional cytokines
XX Sequence 18 AA;

Query Match 59.8%; Score 58; DB 4; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035; Mismatches 3; Indels 0; Gaps 0;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQWLKSR 14
:|||||:|||||
Db 1 SIEGPTLRWLTSR 14

RESULT 39
ABB72912
ID ABB72912 standard; peptide; 18 AA.
XX ABB72912;
AC ABB72912;
XX 05-APR-2002 (first entry)
DT TPO mimetic peptide SRQ ID NO:82.
DE XX
XX Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO200183525-A2.
PN
XX 08-NOV-2001.
PD
XX 02-MAY-2001; 2001WO-US014310.
PF
XX 03-MAY-2000; 2000US-00563286.
PR
XX (AMGE-) AMGEN INC.
PA
XX Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
PI WPI; 2002-130313/17.
XX
XX Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
XX Claim 39; Page 44; 176pp; English.
XX
XX The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, anorectic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic

CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 5; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035; 1; Indels 0; Gaps 0;
Matches 10; Conservative 3; Mismatches 0;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLRWLTSR 14
:|:|||||:|:|

RESULT 40
ADJ73064
ID ADJ73064 standard; peptide; 18 AA.
XX
AC ADJ73064;
XX
DT 06-MAY-2004 (first entry)
XX
DE TPO mimetic peptide sequence SeqID 518.
XX
XX mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW TPO.
XX
XX Synthetic.
OS
PN WO2003084477-A2.
XX
XX 16-OCT-2003.
XX
XX 24-MAR-2003; 2003WO-US009139.
XX
XX 29-MAR-2002; 2002US-0368791P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Scallion BJ, Ghayeb J;
PI WPI; 2003-804237/75.
XX
XX New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
XX Disclosure; SEQ ID NO 518; 97pp; English.

CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating and/or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a TPO mimetic peptide sequence used to make a
CC mimetibody of the invention.

XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 7; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035; 1; Indels 0; Gaps 0;
Matches 10; Conservative 3; Mismatches 1;

QY 1 TIKGPTLRQWLKSR 14
Db 1 SIEGPTLRWLTSR 14
:|:|||||:|:|

RESULT 41
ADJ52699
ID ADJ52699 standard; peptide; 18 AA.
XX
AC ADJ52699;
XX
DT 06-MAY-2004 (first entry)
XX
DE CHI deleted mimetibody-related peptide SeqID518.
XX
XX CHI deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
XX Unidentified.
OS
OS Synthetic.
XX
PN WO2004002417-A2.
XX
XX 08-JAN-2004.
XX
XX 27-JUN-2003; 2003WO-US020347.
XX
XX 28-JUN-2002; 2002US-0392431P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Nesspor TC;
PI Kutoloski KA;
XX
XX WPI; 2004-082870/08.
XX
XX New CHI-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
XX Claim 2; SEQ ID NO 518; 129pp; English.

CC This invention relates to CHI deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CHI-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.

XX
SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 8; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.

XX

SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 8; Length 18;
Best Local Similarity 71.4%; Pred. No. 0.035;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TIKGPTLRQLKSR 14
:|:||||||:|
DB 1 SIEGPTLREWLTSR 14

RESULT 43

ID ADV44468 standard; peptide; 18 AA.
XX
AC ADV44468;
DT 10-MAR-2005 (first entry)
XX Agonist TPO mimetic peptide seqid 152.
DE
XX Anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
KW transplant rejection; prophylaxis; myeloproliferative disorder;
KW hematological disease; TPO mimetic peptide.
XX
OS Unidentified.
XX
XX WO2004108078-A2.
XX
PD 16-DEC-2004.
XX
PF 26-MAY-2004; 2004WO-US016574.
XX
PR 02-JUN-2003; 2003US-00452590.
XX
PA (ALEX-) ALEXION PHARM INC.
XX
PI Bowdish KS, Frederickson S, Renshaw M, Orenca C;
XX WPI; 2005-031588/03.
DR
XX New immunoglobulin molecule comprises a region where amino acid residues
PT corresponding to a portion of complementarity determining region (CDR) is
PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
PT patients.
XX
PS Example 8; SEQ ID NO 152; 139pp; English.
XX
CC The invention describes an immunoglobulin molecule or its fragment
CC comprising: a region where amino acid residues corresponding to at least
CC a portion of two CDRA are replaced with a peptide mimetic selected from
CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced by a peptide
CC mimetic including SEQ ID NO. 126 (not defined in the specification),
CC where X at each occurrence represents any amino acid. Also described are:
CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
CC expression vector comprising the nucleic acid of (1); a host cell
CC transformed with the expression vector of (2); producing an
CC immunoglobulin molecule or its fragment; and a composition comprising the
CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
CC disclosed are: engineering immunoglobulin molecules or fragments;
CC creation of a library of monoclonal antibodies; stimulating
CC proliferation, differentiation, or growth of megakaryocytes; and
CC activating a homodimeric receptor protein. The immunoglobulin molecules
CC are useful for treating HIV-infected patients, patients undergoing
CC chemotherapy, bone marrow transplant patients, stem cell transplant
CC patients, or patients suffering from myeloproliferative disorders. This
CC is the amino acid sequence of an agonist TPO mimetic peptide used to

QY 1 TIKGPTLRQLKSR 14
:|:||||||:|
DB 1 SIEGPTLREWLTSR 14

RESULT 42

ID ADV51660 standard; peptide; 18 AA.
XX
AC ADV51660;
XX
DT 06-MAY-2004 (first entry)
XX
XX CH1 deleted mimetibody-related peptide SeqID518.
XX
XX CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cystostatic; antiinflammatory; neuroleptic;
KW ophthalmologic; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematological disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PP 30-JUN-2003; 2003WO-US020495.
XX
XX 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
XX (CENZ) CENTOCOR INC.
XX
XX Heavner GA, Knight DM, Ghayeb J, Scallion BJ, Neespor TC;
PI KutoLoski KA;
XX WPI; 2004-082872/08.
DR
XX New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
XX Claim 15; SEQ ID NO 518; 123pp; English.
PS
XX This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytotstatic,
CC antiinflammatory, neuroleptic, ophthalmologic, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,

CC replace a CDR in a rationally designed antibody.

XX Sequence 18 AA;

SQ Query Match 59.8%; Score 58; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.035;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHPT 17
|:||||| : : :
Db 3 IEGPTLRQWLAARANS 18

RESULT 44

AE12939

ID AEB12939 standard; peptide; 18 AA.

XX

AC AEB12939;

XX

DT 08-SEP-2005 (first entry)

XX

DE TPO mimetic peptide for VH CDR2 grafting.

XX

XX TPO; Thrombopoietin; protein therapy; antibody engineering;
KW hemipoliosis; immunotherapy; Cardiant; Antidiabetic; Anorectic;
KW cardiac failure; diabetes; obesity.

XX Synthetic.

XX WO2005060642-A2.

PN

XX

PD 07-JUL-2005.

XX

XX 15-DEC-2004; 2004WO-US041946.

XX

XX 15-DEC-2003; 2003US-00737290.

PR

XX (ALEX-) ALEXION PHARM INC.

PA

XX Bowdish KS, Frederickson S, Renshaw M, Orenica C;

PI

XX WPI; 2005-479402/48.

XX

XX New immunoglobulin molecule comprising a region where amino acid residues
PT corresponding to at least a portion of a complementarity determining
PT region is replaced with a peptide, for treating congestive heart failure,
PT diabetes or obesity.

XX

XX Example 8; SEQ ID NO 152; 152pp; English.

XX

XX The invention relates to an immunoglobulin (Ig) molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a complementarity determining regions (CDR) is replaced with a
CC peptide selected from human brain natriuretic protein (hBNP), hBNP
CC mimetics, glycogen phosphorylase (GLP)-1, GLP-1 mimetics, GLP-2, GLP-2
CC mimetics, exendin, exendin mimetics, glucagons, glucagon mimetics and
CC PACAP-38. Also included are a nucleic acid encoding the immunoglobulin
CC molecule, an expression vector comprising the nucleic acid, a host cell
CC transformed with the expression vector, producing an immunoglobulin
CC molecule (or its fragment, comprising culturing the host cell under
CC conditions suitable for expression of the immunoglobulin or its
CC fragment), a composition comprising an immunoglobulin (or its fragment)
CC and a pharmaceutically acceptable carrier, treating congestive heart
CC failure (comprising administering to the subject an immunoglobulin
CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a peptide
CC selected from hBNP and hBNP mimetics), treating diabetes or obesity
CC (comprising administering to a subject an immunoglobulin molecule or its
CC fragment comprising a region where amino acid residues corresponding to
CC at least a portion of a CDR is replaced with a peptide selected from GLP-
CC 1, GLP-1 mimetics, GLP-2, GLP-2 mimetics, exendin, exendin mimetics,
CC glucagons, glucagon mimetics and PACAP-38), preserving/improving beta-
CC cell function (comprising administering to a subject an immunoglobulin

CC molecule or fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with GLP-1,
CC inducing endothelial-dependent relaxation of precontracted pulmonary
CC artery rings (comprising administering to a subject an immunoglobulin
CC molecule or fragment thereof comprising a region where amino acid
CC residues corresponding to at least a portion of a CDR is replaced with
CC GLP-1) and administering to a subject an immunoglobulin molecule or its
CC fragment (comprising a region where amino acid residues corresponding to
CC at least a portion of a complementarity determining regions (CDR) is
CC replaced with a thiazolidinedione derivative), regulating adiponectin
CC expression (comprising administering to a subject an immunoglobulin
CC molecule or its fragment comprising a region where amino acid residues
CC corresponding to at least a portion of a CDR is replaced with a
CC thiazolidinedione derivative). The immunoglobulin is an anti-tetanus
CC toxoid antibody (TT) where the heavy chain CDR2 and/or CDR3 are fully or
CC partially replaced with a peptide listed above or (as described in the
CC examples) a Thrombopoietin (TPO) mimetic, erythropoietin (EPO) mimetic
CC or ANP (atrial natriuretic peptide). The molecule, composition and
CC methods are useful for treating congestive heart failure, diabetes or
CC obesity. The present sequence is a TPO mimetic peptide (full length, core
CC or flanked by random amino acids) for inclusion in an immunoglobulin of
CC the invention.

XX

SQ Sequence 18 AA;

Query Match 59.8%; Score 58; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.035;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHPT 17
|:||||| : : :
Db 3 IEGPTLRQWLAARANS 18

RESULT 45

ADQ16705

ID ADQ16705 standard; protein; 128 AA.

XX

AC ADQ16705;

XX

DT 09-SEP-2004 (first entry)

XX

DE Modified immunoglobulin clone 116 HC variable region SEQ ID NO:125.

XX

XX immunoglobulin; complementarity determining region; CDR; peptide mimetic;
KW erythropoietin; EPO; thrombopoietin; TPO; immunosuppressive;
KW immunotherapy; thrombocytopenia.

XX Synthetic.

OS

XX WO2004050017-A2.

PN

XX 17-JUN-2004.

PD

XX 17-NOV-2003; 2003WO-US036894.

PF

XX 02-DEC-2002; 2002US-00307724.

PR

XX (ALEX-) ALEXION PHARM INC.

PA

XX Bowdish KS, Frederickson S, Renshaw M;

PI

XX WPI; 2004-460973/43.

XX

XX New immunoglobulin molecule comprising a region, where two
PT complementarity determining regions (CDRs) are replaced with EPO mimetic
PT or a TPO mimetic, useful for treating thrombocytopenia.

XX

XX Claim 9; SEQ ID NO 125; 107pp; English.

PS

XX The invention relates to a novel immunoglobulin molecule or its fragment
CC comprising a region where amino acid residues corresponding to at least a
CC portion of a two complementarity determining regions (CDRs) are replaced

CC with a peptide mimetic selected from an erythropoietin (EPO) mimetic and
 CC a thrombopoietin (TPO) mimetic. An immunoglobulin molecule of the
 CC invention has immunosuppressive activity, and may have a use in
 CC immunotherapy. The immunoglobulin molecule is useful for diagnosing or
 CC treating thrombocytopenia as a result of chemotherapy, bone marrow
 CC transplantation, or chronic diseases such as idiopathic thrombocytopenia.
 CC The present sequence represents immunoglobulin clone 116 heavy chain
 CC variable region.
 XX
 XX
 SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 8; Length 128;
 Best Local Similarity 62.5%; Pred. No. 0.34;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHT 17
 |:|||||: : :
 Db 52 IEGPTLRQWLARANS 67

RESULT 46
 ADV44466
 ID ADV44466 standard; protein; 128 AA.

XX AC ADV44466;

XX DT 10-MAR-2005 (first entry)

XX DE Anti-tetanus toxoid Fab pAX116 heavy chain variant PRL5-116 13F2.

XX KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
 KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
 KW transplant rejection; prophylaxis; myeloproliferative disorder;
 KW hematological disease; pAX116.

XX OS Homo sapiens.
 OS Synthetic.
 OS Unidentified.

XX PN WO2004108078-A2.

XX PD 16-DEC-2004.

XX PP 26-MAY-2004; 2004WO-US016574.

XX PR 02-JUN-2003; 2003US-00452590.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Frederickson S, Renshaw M, Orenca C;

XX DR WPI; 2005-031588/03.

XX PT New immunoglobulin molecule comprises a region where amino acid residues
 PT corresponding to a portion of complementarity determining region (CDR) is
 PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
 PT patients.

XX PS Claim 148; SEQ ID NO 150; 139pp; English.

XX CC The invention describes an immunoglobulin molecule or its fragment
 CC comprising: a region where amino acid residues corresponding to at least
 CC a portion of two CDRs are replaced with a peptide mimetic selected from
 CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced by a peptide
 CC mimetic including SEQ ID NO. 126 (not defined in the specification),
 CC where X at each occurrence represents any amino acid. Also described are:
 CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
 CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an
 CC immunoglobulin molecule or its fragment; and a composition comprising the
 CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
 CC disclosed are: engineering immunoglobulin molecules or fragments;

CC creation of a library of monoclonal antibodies; stimulating
 CC proliferation, differentiation, or growth of megakaryocytes; and
 CC activating a homodimeric receptor protein. The immunoglobulin molecules
 CC are useful for treating HIV-infected patients, patients undergoing
 CC chemotherapy, bone marrow transplant patients, stem cell transplant
 CC patients, or patients suffering from myeloproliferative disorders. This
 CC is the amino acid sequence of a variant of the anti-tetanus toxoid
 CC antibody pAX116 heavy chain, CDR2 of which has been partially replaced
 CC with residues of a TPO-mimetic peptide.
 XX
 XX
 SQ Sequence 128 AA;

Query Match 59.8%; Score 58; DB 9; Length 128;
 Best Local Similarity 62.5%; Pred. No. 0.34;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 IKGPTLRQWLKSRHT 17
 |:|||||: : :
 Db 52 IEGPTLRQWLARANS 67

RESULT 47
 ADV44463

ID ADV44463 standard; protein; 128 AA.

XX AC ADV44463;

XX DT 10-MAR-2005 (first entry)

XX DE Anti-tetanus toxoid Fab pAX116 heavy chain variant PRL5-116.

XX KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
 KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
 KW transplant rejection; prophylaxis; myeloproliferative disorder;
 KW hematological disease; pAX116.

XX OS Homo sapiens.
 OS Synthetic.
 OS Unidentified.

XX PN WO2004108078-A2.

XX PD 16-DEC-2004.

XX PP 26-MAY-2004; 2004WO-US016574.

XX PR 02-JUN-2003; 2003US-00452590.

XX PA (ALEX-) ALEXION PHARM INC.

XX PI Bowdish KS, Frederickson S, Renshaw M, Orenca C;

XX DR WPI; 2005-031588/03.

XX PT New immunoglobulin molecule comprises a region where amino acid residues
 PT corresponding to a portion of complementarity determining region (CDR) is
 PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
 PT patients.

XX PS Example 11; SEQ ID NO 147; 139pp; English.

XX CC The invention describes an immunoglobulin molecule or its fragment
 CC comprising: a region where amino acid residues corresponding to at least
 CC a portion of two CDRs are replaced with a peptide mimetic selected from
 CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced by a peptide
 CC mimetic including SEQ ID NO. 126 (not defined in the specification),
 CC where X at each occurrence represents any amino acid. Also described are:
 CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
 CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an
 CC immunoglobulin molecule or its fragment; and a composition comprising the
 CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
 CC disclosed are: engineering immunoglobulin molecules or fragments;

CC disclosed are: engineering immunoglobulin molecules or fragments;
 CC creation of a library of monoclonal antibodies; stimulating
 CC proliferation, differentiation, or growth of megakaryocytes; and
 CC activating a homodimeric receptor protein. The immunoglobulin molecules
 CC are useful for treating HIV-infected patients, patients undergoing
 CC chemotherapy, bone marrow transplant patients, stem cell transplant
 CC patients, or patients suffering from myeloproliferative disorders. This
 CC is the amino acid sequence of a variant of the anti-tetanus toxoid
 CC antibody pAX116 heavy chain, CDR2 of which has been partially replaced
 CC with residues of a TPO-mimetic peptide.
 XX
 SQ Sequence 128 AA;
 Query Match 59.8%; Score 58; DB 9; Length 128;
 Best Local Similarity 62.5%; Pred. No. 0.34;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 IKGPTLRQWLKSRHPT 17
 :||||||| : :
 Db 52 IEGPTLRQWLARANS 67
 RESULT 48
 ADV44467
 ID ADV44467 standard; protein; 128 AA.
 XX
 AC ADV44467;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE Anti-tetanus toxoid Fab pAX116 heavy chain variant pRL5-116 XX12.
 XX
 KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
 KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
 KW transplant rejection; prophylaxis; myeloproliferative disorder;
 KW hematological disease; pAX116.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS Unidentified.
 XX
 PN WO2004108078-A2.
 XX
 PD 16-DEC-2004.
 XX
 PF 26-MAY-2004; 2004WO-US016574.
 XX
 PR 02-JUN-2003; 2003US-00452590.
 XX
 PA (ALEX-) ALEXION PHARM INC.
 XX
 PI Bowdish KS, Frederickson S, Renshaw M, Orenia C;
 XX
 DR WPI; 2005-031588/03.
 XX
 PT New immunoglobulin molecule comprises a region where amino acid residues
 PT corresponding to a portion of complementarity determining region (CDR) is
 PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
 PT patients.
 XX
 PS Claim 148; SEQ ID NO 151; 139pp; English.
 XX
 CC The invention describes an immunoglobulin molecule or its fragment
 CC comprising: a region where amino acid residues corresponding to at least
 CC a portion of two CDRs are replaced with a peptide mimetic selected from
 CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced by a peptide
 CC mimetic including SEQ ID NO. 126 (not defined in the specification),
 CC where X at each occurrence represents any amino acid. Also described are:
 CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
 CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an
 CC immunoglobulin molecule or its fragment; and a composition comprising the

CC immunoglobulin or its fragment and a pharmaceutical carrier. Also
 CC disclosed are: engineering immunoglobulin molecules or fragments;
 CC creation of a library of monoclonal antibodies; stimulating
 CC proliferation, differentiation, or growth of megakaryocytes; and
 CC activating a homodimeric receptor protein. The immunoglobulin molecules
 CC are useful for treating HIV-infected patients, patients undergoing
 CC chemotherapy, bone marrow transplant patients, stem cell transplant
 CC patients, or patients suffering from myeloproliferative disorders. This
 CC is the amino acid sequence of a variant of the anti-tetanus toxoid
 CC antibody pAX116 heavy chain, CDR2 of which has been partially replaced
 CC with residues of a TPO-mimetic peptide.
 XX
 SQ Sequence 128 AA;
 Query Match 59.8%; Score 58; DB 9; Length 128;
 Best Local Similarity 62.5%; Pred. No. 0.34;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 IKGPTLRQWLKSRHPT 17
 :||||||| : :
 Db 52 IEGPTLRQWLARANS 67
 RESULT 49
 ADV44465
 ID ADV44465 standard; protein; 128 AA.
 XX
 AC ADV44465;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE Anti-tetanus toxoid Fab pAX116 heavy chain variant pRL5-116 10B12.
 XX
 KW anti-HIV; cytostatic; gene therapy; antibody engineering; diagnosis;
 KW HIV-infection; anti-HIV; chemotherapy; bone marrow transplantation;
 KW transplant rejection; prophylaxis; myeloproliferative disorder;
 KW hematological disease; pAX116.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS Unidentified.
 XX
 PN WO2004108078-A2.
 XX
 PD 16-DEC-2004.
 XX
 PF 26-MAY-2004; 2004WO-US016574.
 XX
 PR 02-JUN-2003; 2003US-00452590.
 XX
 PA (ALEX-) ALEXION PHARM INC.
 XX
 PI Bowdish KS, Frederickson S, Renshaw M, Orenia C;
 XX
 DR WPI; 2005-031588/03.
 XX
 PT New immunoglobulin molecule comprises a region where amino acid residues
 PT corresponding to a portion of complementarity determining region (CDR) is
 PT replaced with a peptide mimetic, useful for treating, e.g. HIV-infected
 PT patients.
 XX
 PS Claim 148; SEQ ID NO 149; 139pp; English.
 XX
 CC The invention describes an immunoglobulin molecule or its fragment
 CC comprising: a region where amino acid residues corresponding to at least
 CC a portion of two CDRs are replaced with a peptide mimetic selected from
 CC an EPO mimetic or a TPO mimetic; or a region where amino acid residues
 CC corresponding to at least a portion of a CDR is replaced by a peptide
 CC mimetic including SEQ ID NO. 126 (not defined in the specification),
 CC where X at each occurrence represents any amino acid. Also described are:
 CC a nucleic acid encoding an immunoglobulin molecule or its fragment; an
 CC expression vector comprising the nucleic acid of (1); a host cell
 CC transformed with the expression vector of (2); producing an

